ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

CSCS
Swiss National Supercomputing Centre

The DCA++ Story

How new algorithms, new computers, and innovative software design enable petaflop/s scale simulations of high temperature superconductivity

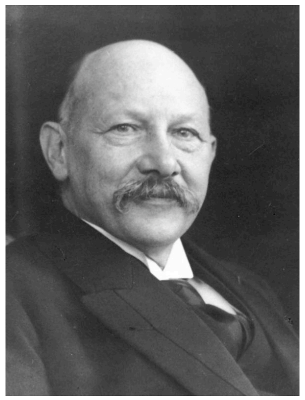
Thomas C. Schulthess
schulthess@cscs.ch

Leadership Computing Facility Seminar,
Oak Ridge National Laboratory 01/30/2009

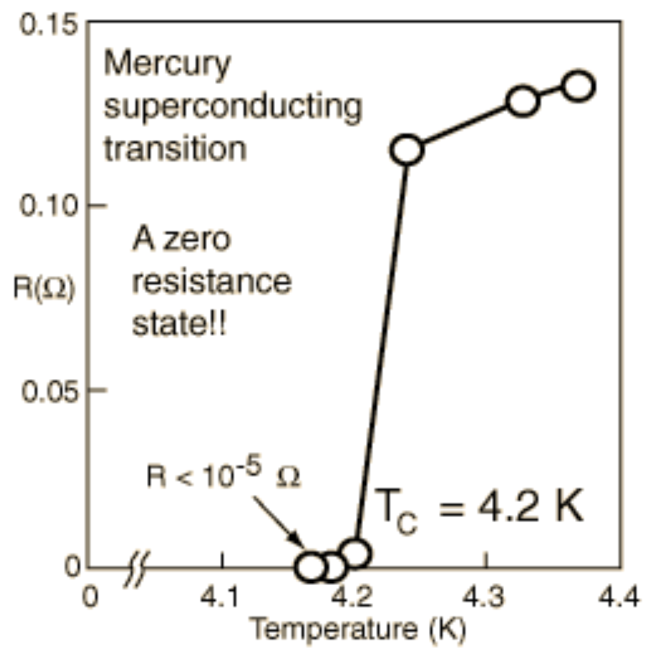
Oak Ridge National Laboratory

Superconductivity: a state of matter with zero electrical resistivity

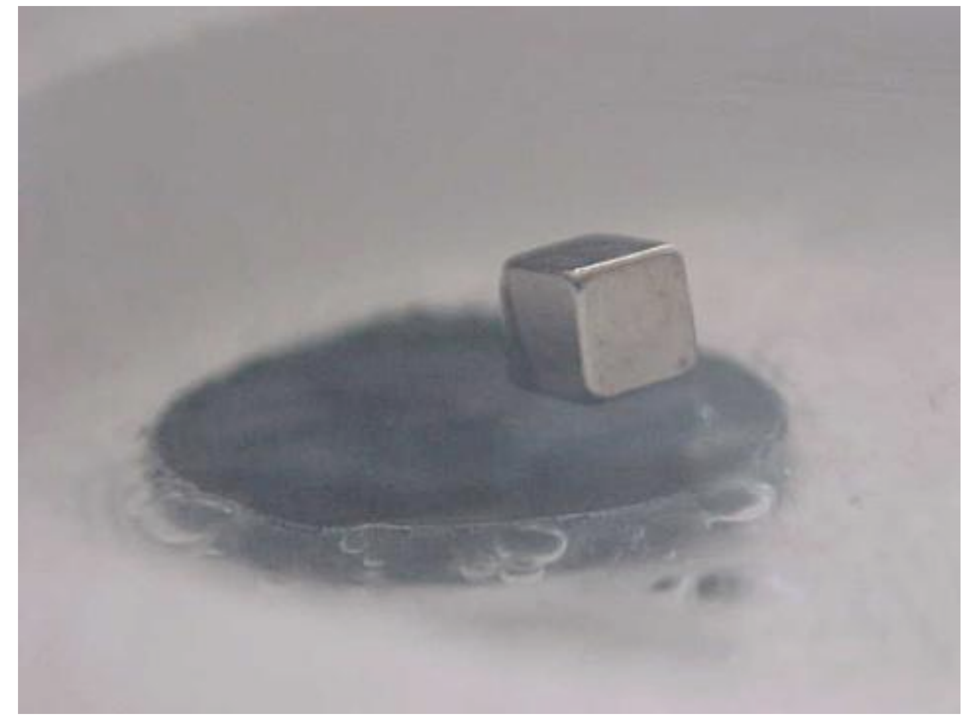
Discovery 1911



Heike Kamerlingh Onnes (1853-1926)



Superconductor repels magnetic field
Meissner and Ochsenfeld, Berlin 1933



Microscopic Theory for Superconductivity 1957

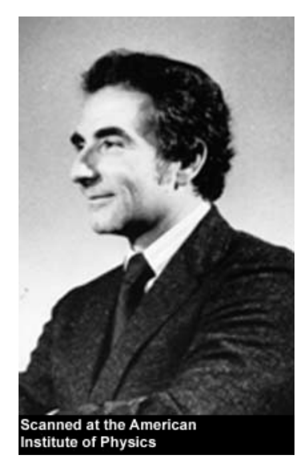
PHYSICAL REVIEW VOLUME 108, NUMBER 5 DECEMBER 1, 1957

Theory of Superconductivity*

J. Bardeen, L. N. Cooper,† and J. R. Schrieffer‡
Department of Physics, University of Illinois, Urbana, Illinois

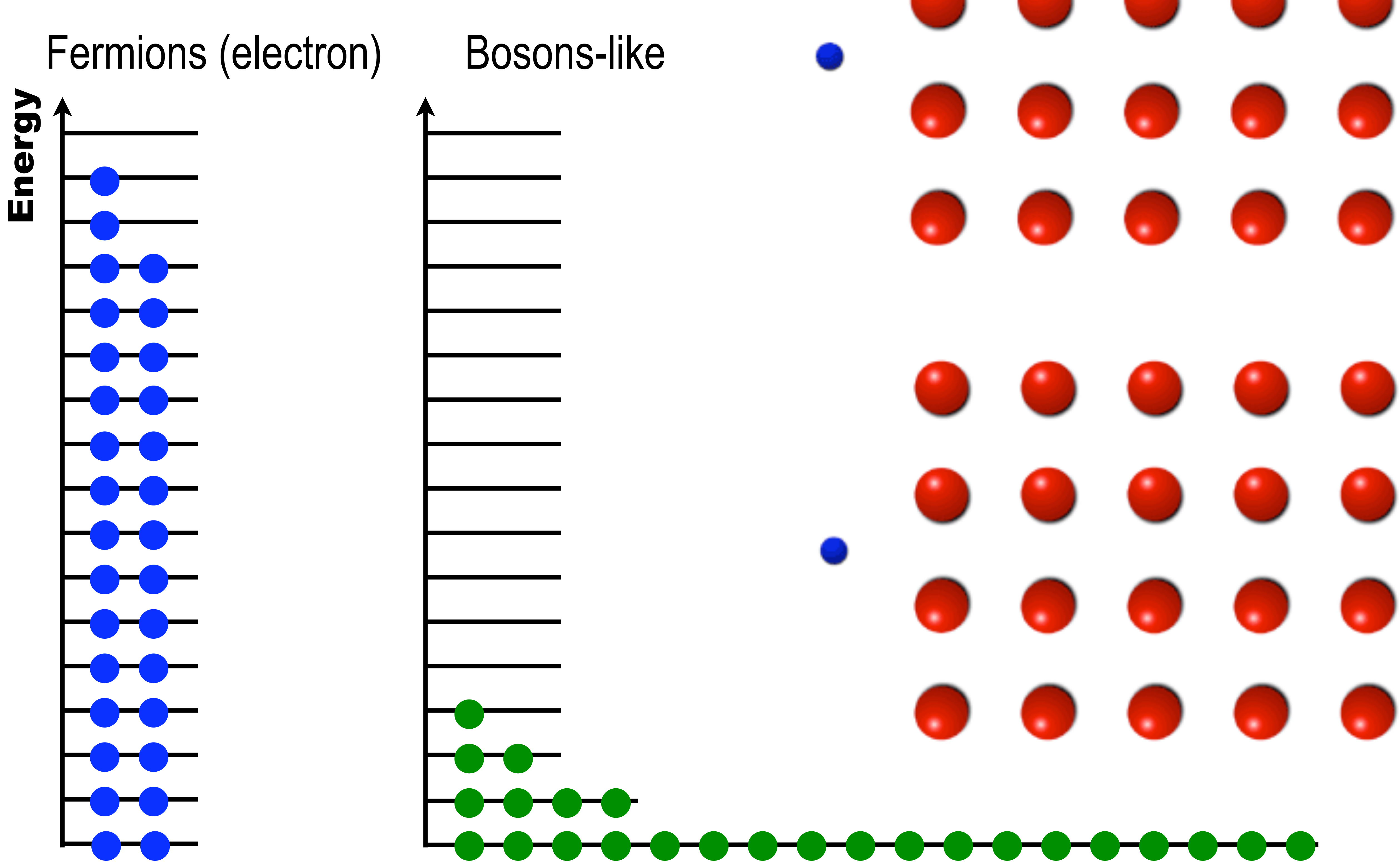
(Received July 8, 1957)

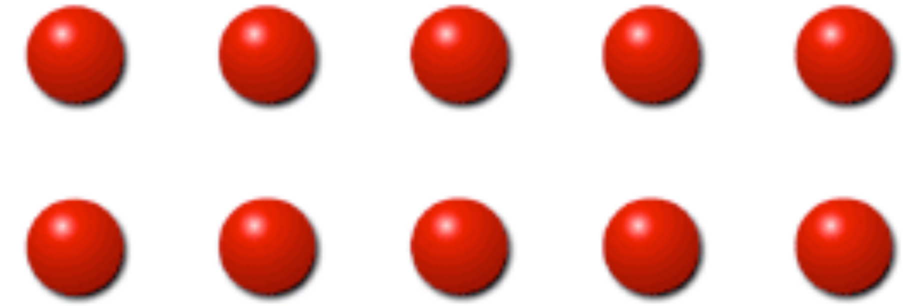
A theory of superconductivity is presented, based on the fact that the interaction between electrons resulting from virtual exchange of phonons is attractive when the energy difference between the electrons states involved is less than the phonon energy, $\hbar\omega$. It is favorable to form a superconducting phase when this attractive interaction dominates the repulsive screened Coulomb interaction. The normal phase is described by the Bloch individual-particle model. The ground state of a superconductor, formed from a linear combination of normal state configurations in which electrons are virtually excited in pairs of opposite spin and momentum, is lower in energy than the normal state by amount proportional to an average $(\hbar\omega)^2$, consistent with the isotope effect. A mutually orthogonal set of excited states in one-to-one correspondence with those of the normal phase is obtained by specifying occupation of certain Bloch states and by using the rest to form a linear combination of virtual pair configurations. The theory yields a second-order phase transition and a Meissner effect in the form suggested by Pippard. Calculated values of specific heats and penetration depths and their temperature variation are in good agreement with experiment. There is an energy gap for individual-particle excitations which decreases from about $3.5kT_c$ at $T=0°$K to zero at T_c. Tables of matrix elements of single-particle operators between the excited-state superconducting wave functions, useful for perturbation expansions and calculations of transition probabilities, are given.



BCS Theory generally accepted in the early 1970s

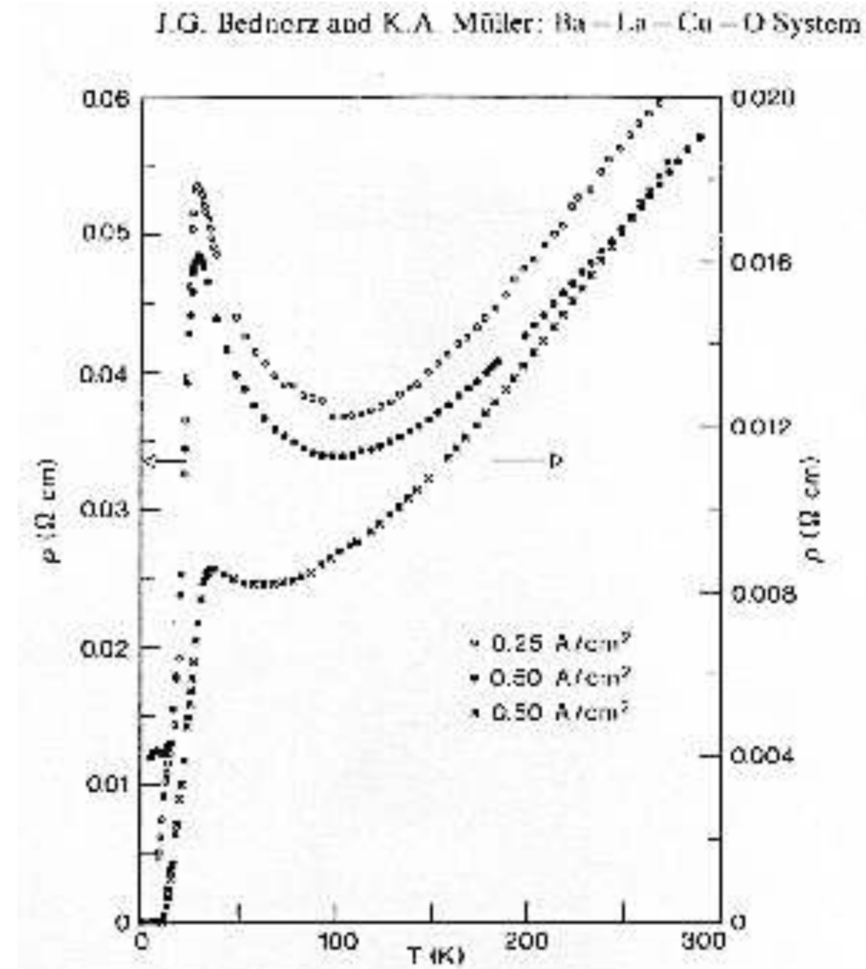
Fermions, Bosons, and Cooper Pairs

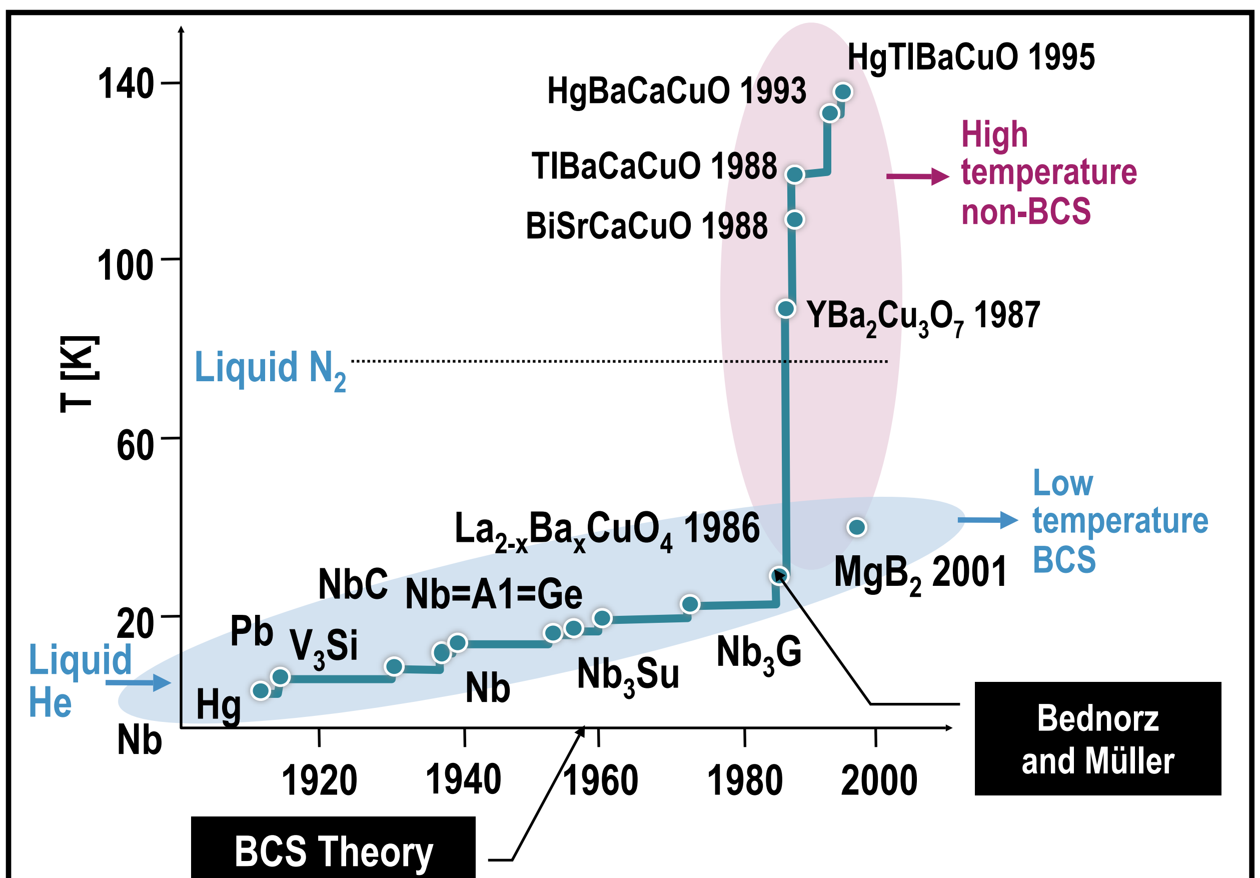




Superconductivity in the cuprates

- High transition temperatures
  - Discovered in 1986 by Bednorz and Müller
- Totally different materials
  - In the normal state
    conventional superconductors are metals
    cuprates are insulators or poor conductors



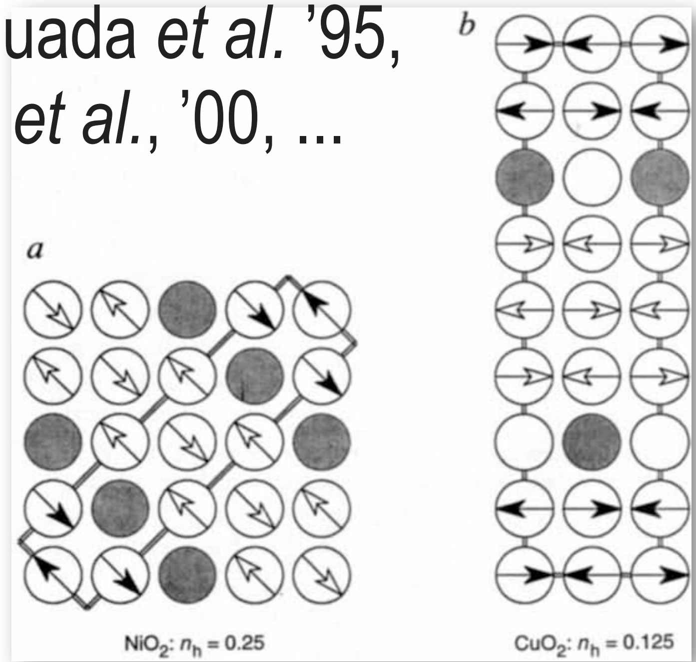


Twenty years later

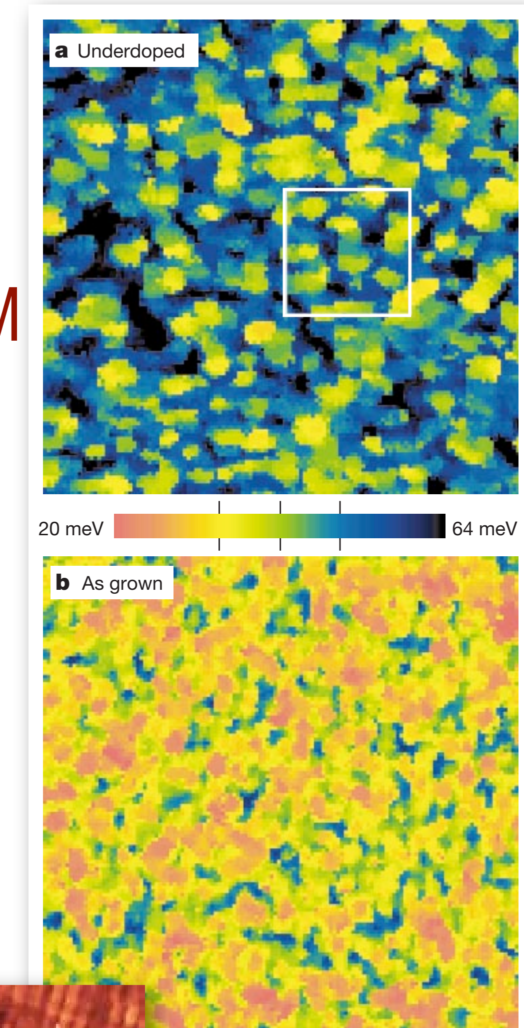
- No predictive power for T_c in known materials
- No predictive power for design of new SC materials
- No explanation for pseudogap phase
- No theory of unusual transport properties
- No controlled solution for proposed effective Hamiltonians
- Only partial consensus on which materials aspects are essential

The role of inhomogeneities

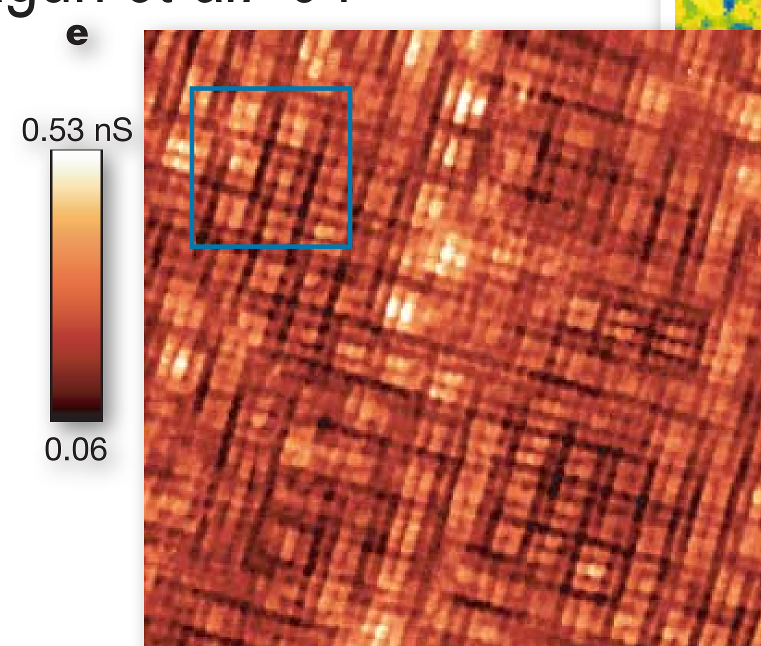
Stripes in neutron scattering:
Tranquada *et al.* '95,
Mook *et al.*, '00, ...



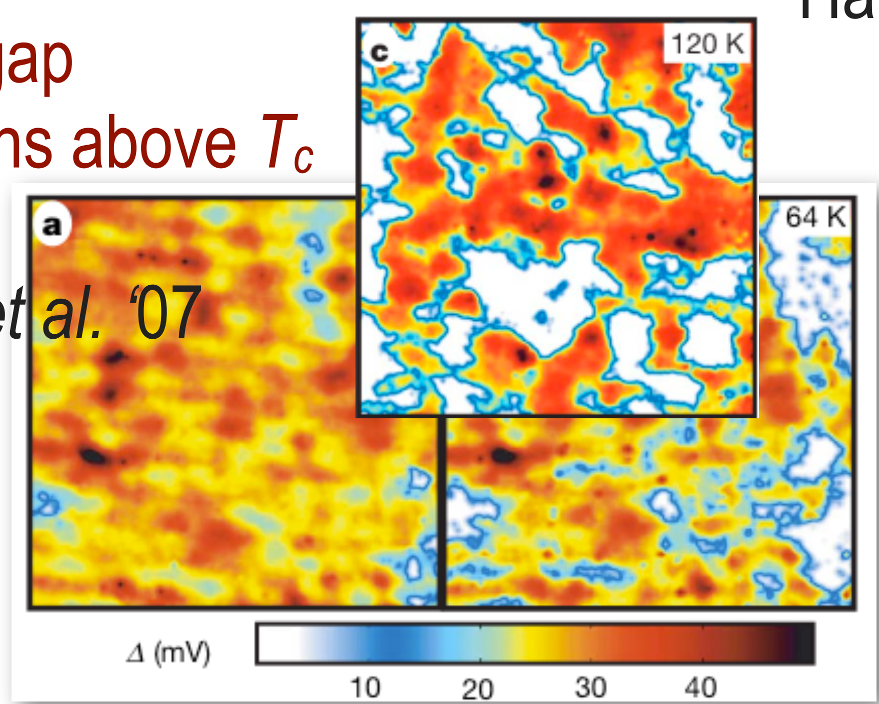
Random SC gap modulations in STM (BSCCO):
Lang *et al.* '02



Charge ordered "checkerboard" state (Na doped cuprates):
Hanaguri *et al.* '04



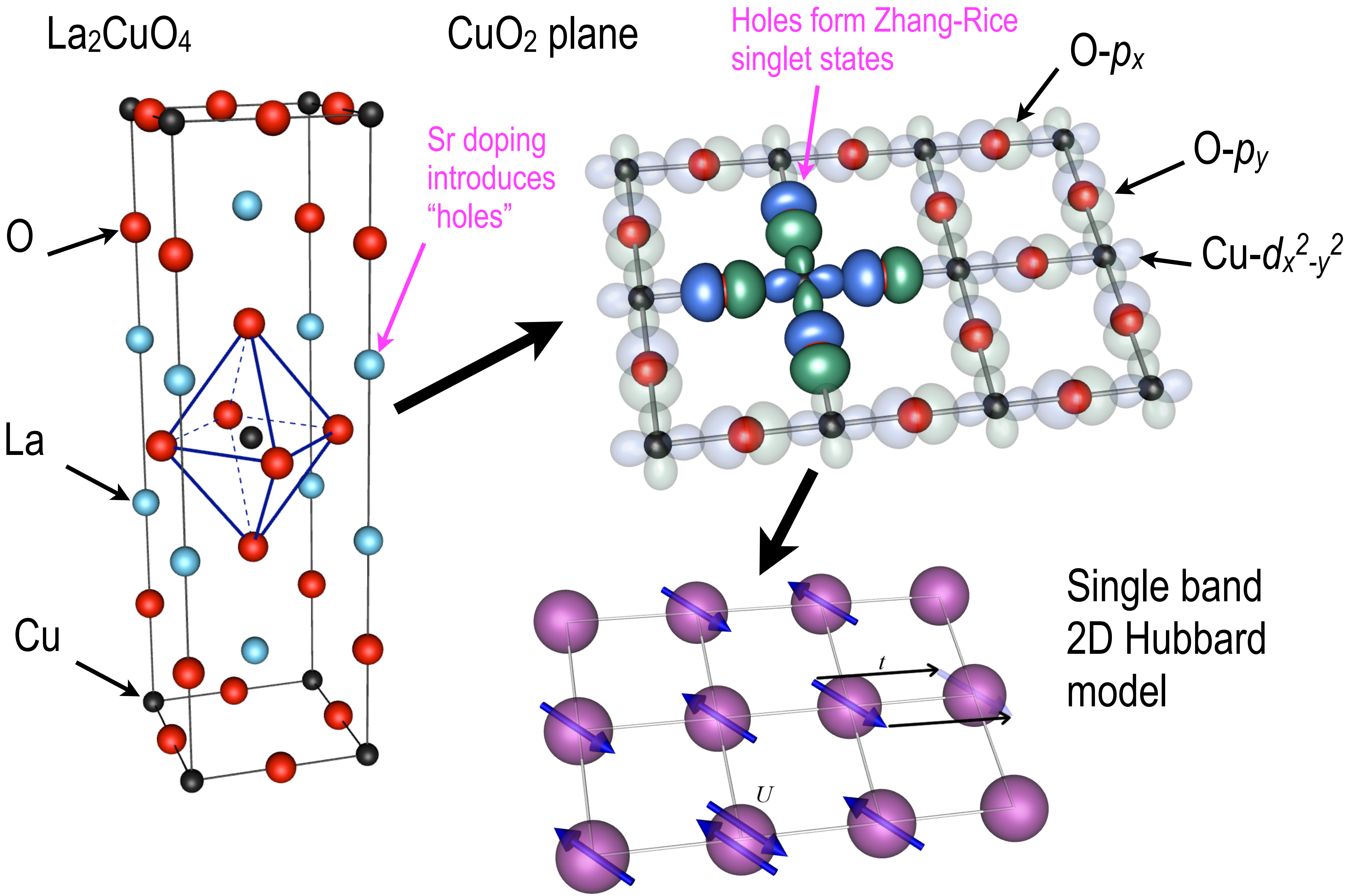
Random gap modulations above T_c (BSCCO):
Gomes *et al.* '07



Outline

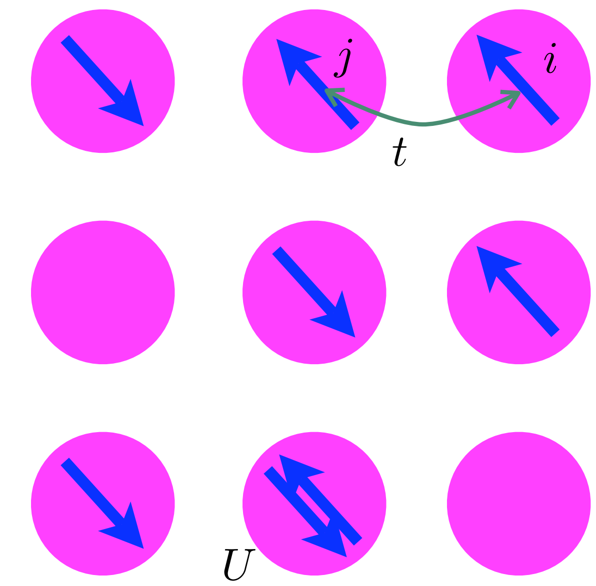
- Brief introduction into superconductivity and the cuprates
- Background: The two dimensional Hubbard model and the DCA/QMC method
- Simulational studies with the DCA/QMC method
- Algorithmic improvements and a method to study effects of disorder an nanoscale inhomogeneities
  - Accelerating Hirsch-Fye QMC with delayed updates
  - Mixed precision and multithreaded implementations (GPU in particular)
  - Disorder averaging and a first study of how disorder affects the superconducting transition temperature
- DCA++, concurrency, scaling, and performance
  - Results for Cray XT4 and first results for a PF/s scale system
- Summary and conclusions

From cuprate materials to the Hubbard model

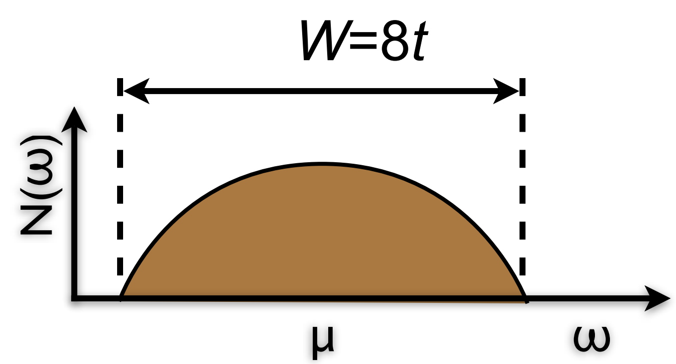


2D Hubbard model and its physics

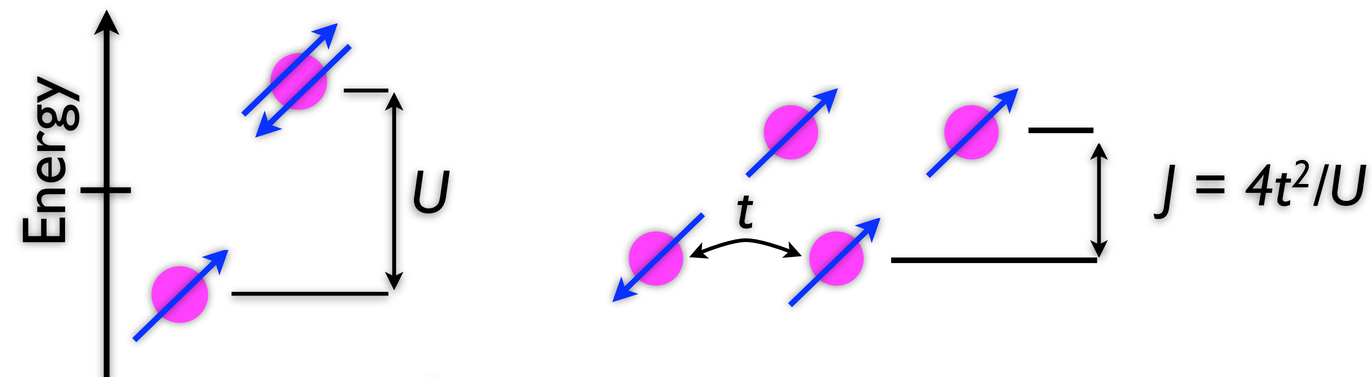
- Hamiltonian $H = -t \sum_{<ij>,\sigma} c^\dagger_{i\sigma} c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}$



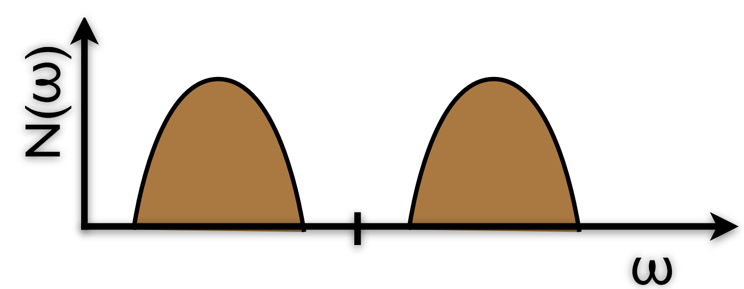
- *U*=0
  - Metallic state with band width *W*=8*t*



- *U*>>8*t*; <*n*>=1 (half-filling)
  - Formation of magnetic moment



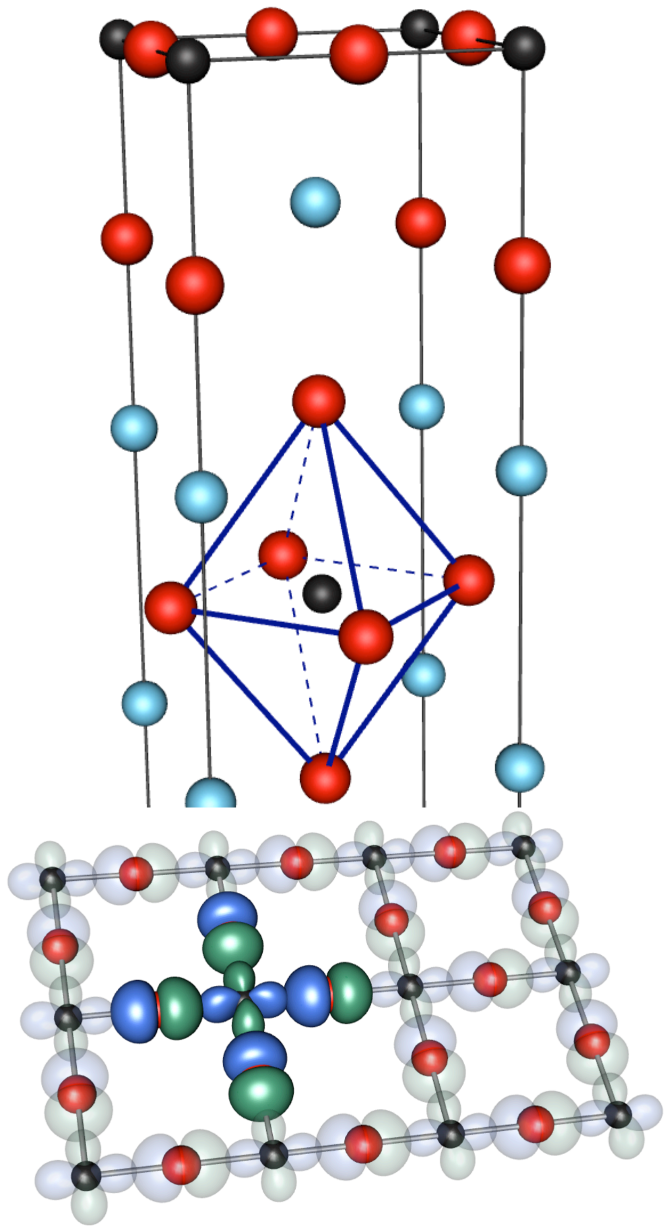
  - Mott insulator, antiferromagnetic ground state



- *U*≈8*t*; filling *δ*=1-<*n*> >0 (parameter range relevant for cuprates)
  - Doped Mott insulator with strong antiferromagnetic correlations

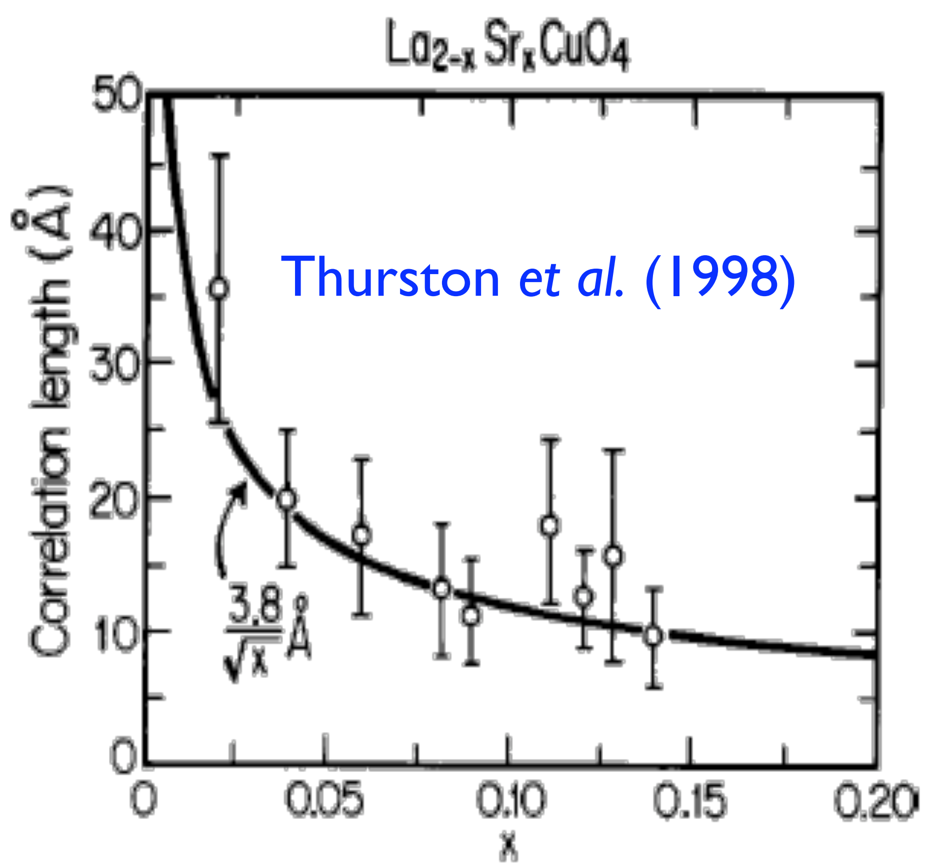
Hamiltonian *H* operates on 4^N dimensional Fock-Space

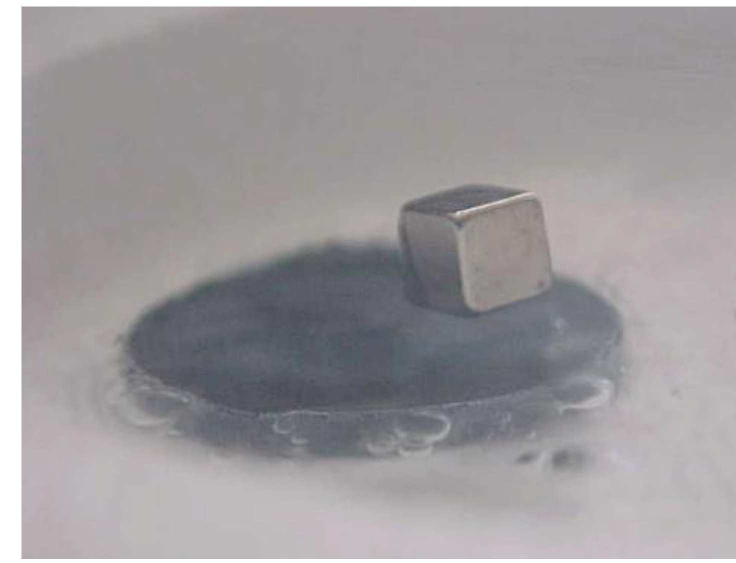
The challenge: a (quantum) multi-scale problem



On-site Coulomb repulsion (~A)

Antiferromagnetic correlations (~nm)





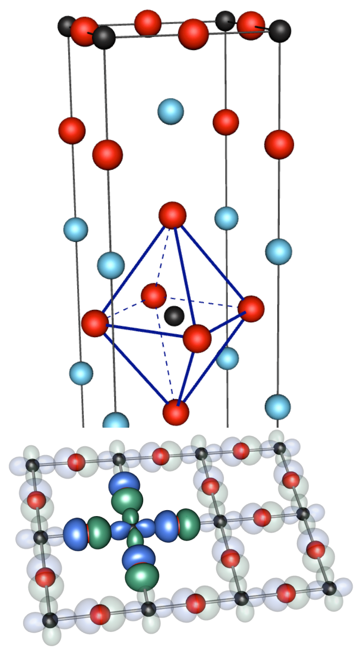
Superconductivity (macroscopic)

$N \sim 10^{23}$

complexity ~ 4^N

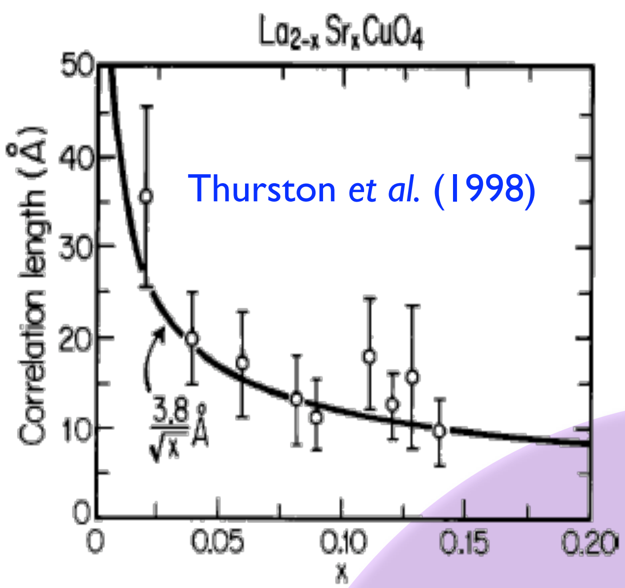
Quantum cluster theories

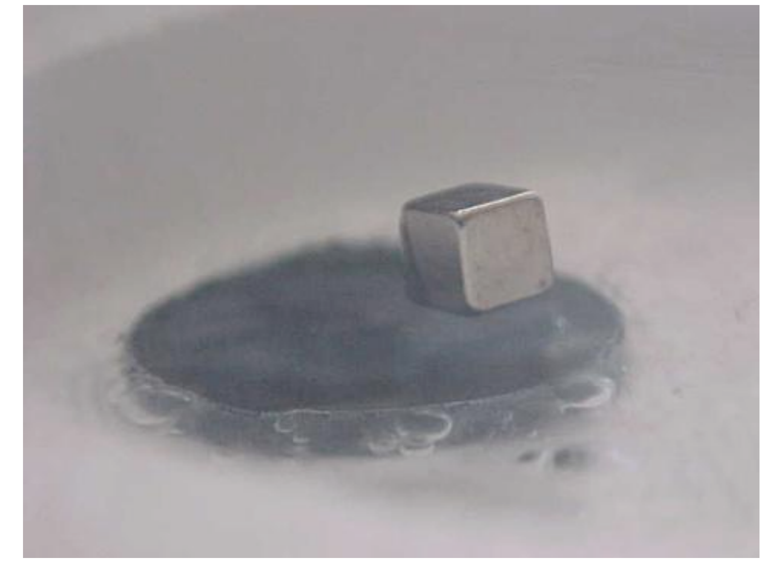
Maier *et al.*, Rev. Mod. Phys. '05



On-site Coulomb repulsion (~A)

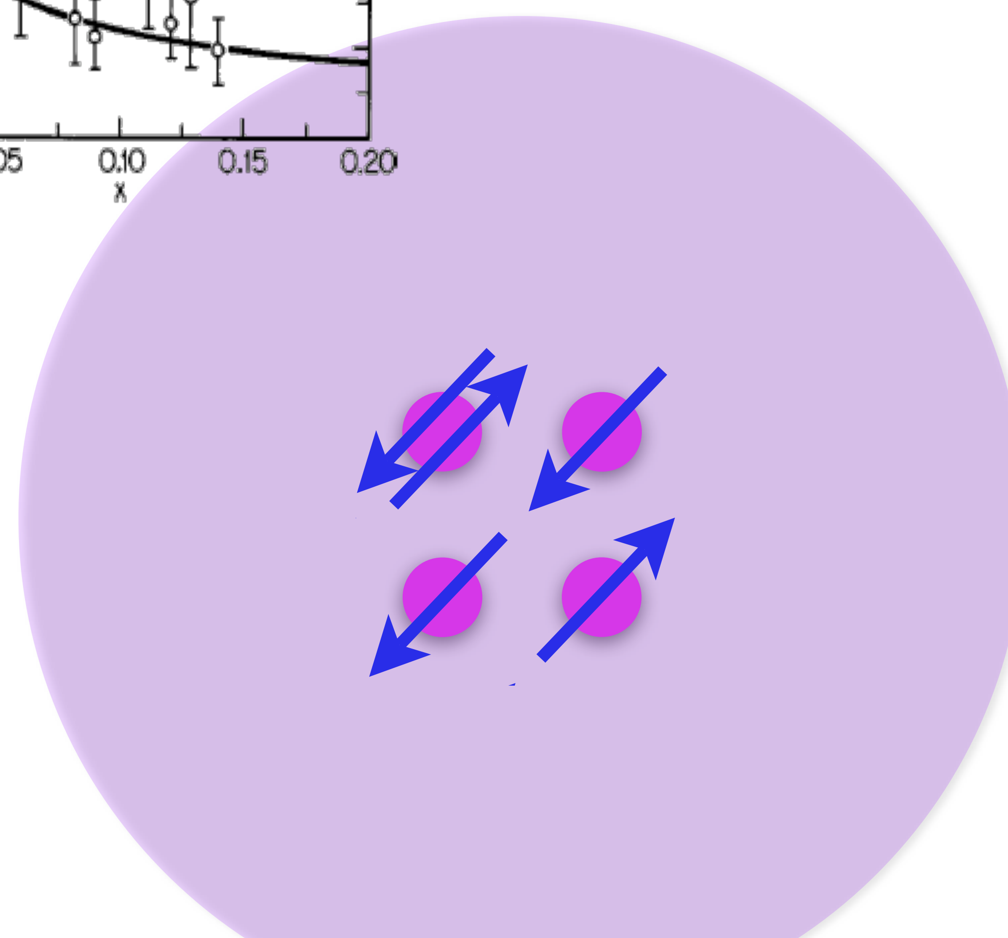
Antiferromagnetic correlations (~nm)





Superconductivity (macroscopic)

Explicitly treat correlations within a localized cluster



Treat macro-scopic scales within mean-field

Coherently embed cluster into effective medium

Green's functions in quantum many-body theory

Noninteracting Hamiltonian &

$$H_0 = \left[-\frac{1}{2}\nabla^2 + V(\vec{r}) \right]$$

Green's function

$$\left[i\frac{\partial}{\partial t} - H_0 \right] G_0(\vec{r}, t, \vec{r}', t') = \delta(\vec{r} - \vec{r}')\delta(t - t)$$

Fourier transform & analytic continuation: $z^{\pm} = \omega \pm i\epsilon$ $\quad G_0^{\pm}(\vec{r}, z) = \left[z^{\pm} - H_0 \right]^{-1}$

Hubbard Hamiltonian

$$H = -t \sum_{<ij>,\sigma} c^\dagger_{i\sigma} c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} \qquad n_{i\sigma} = c^\dagger_{i\sigma} c_{i\sigma}$$

Hide symmetry in algebraic properties of field operators

$$c_{i\sigma} c_{j\sigma'} + c_{j\sigma'} c_{i\sigma} = 0$$

$$c_{i\sigma} c^\dagger_{j\sigma'} + c^\dagger_{j\sigma'} c_{i\sigma} = \delta_{ij}\delta_{\sigma\sigma'}$$

Green's function

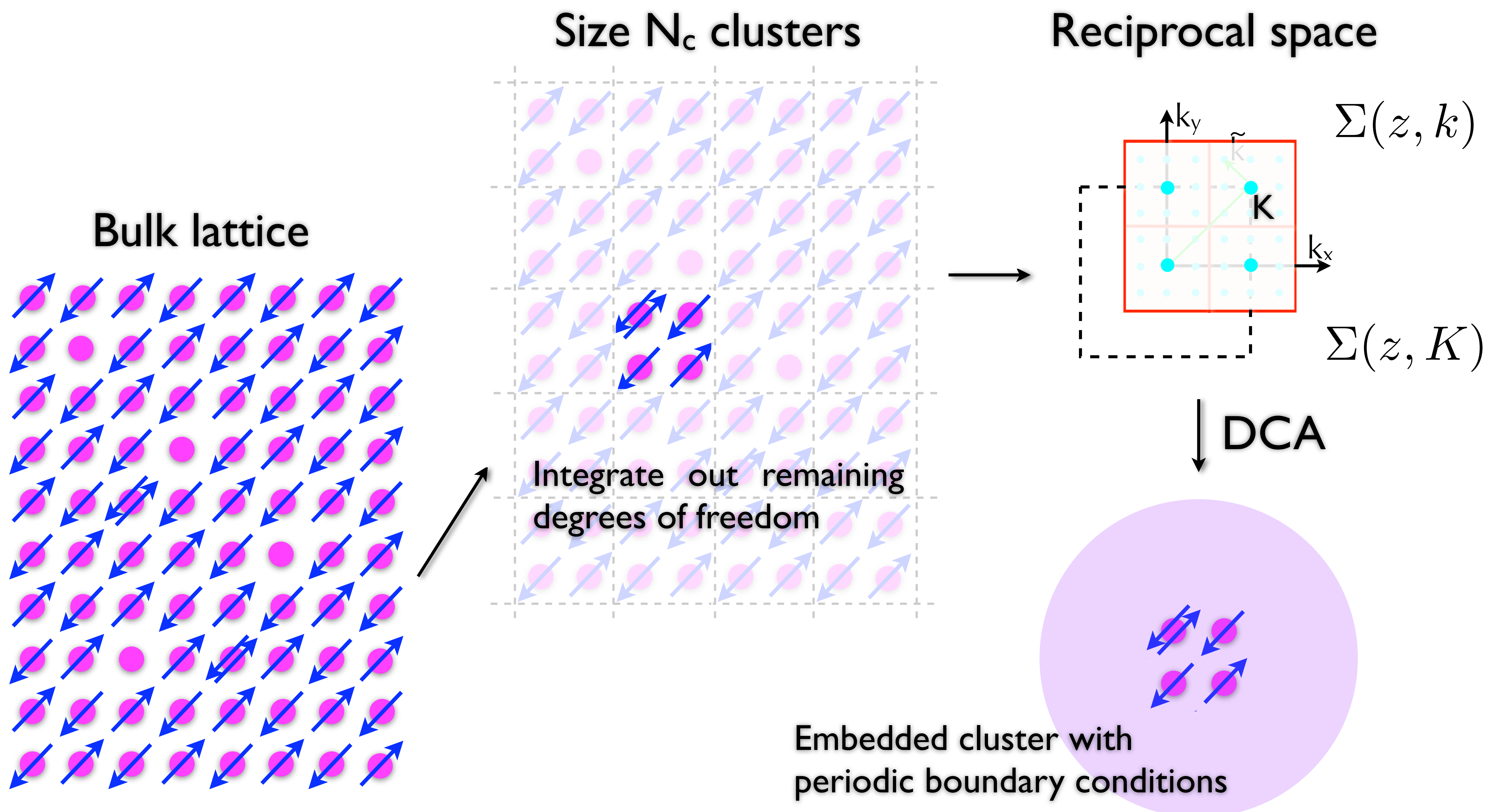
$$G_\sigma(r_i, \tau; r_j, \tau') = -\left\langle \mathcal{T} c_{i\sigma}(\tau) c^\dagger_{j\sigma}(\tau') \right\rangle$$

Spectral representation

$$G_0(k, z) = [z - \epsilon_0(k)]^{-1}$$

$$G(k, z) = [z - \epsilon_0(k) - \Sigma(k, z)]^{-1}$$

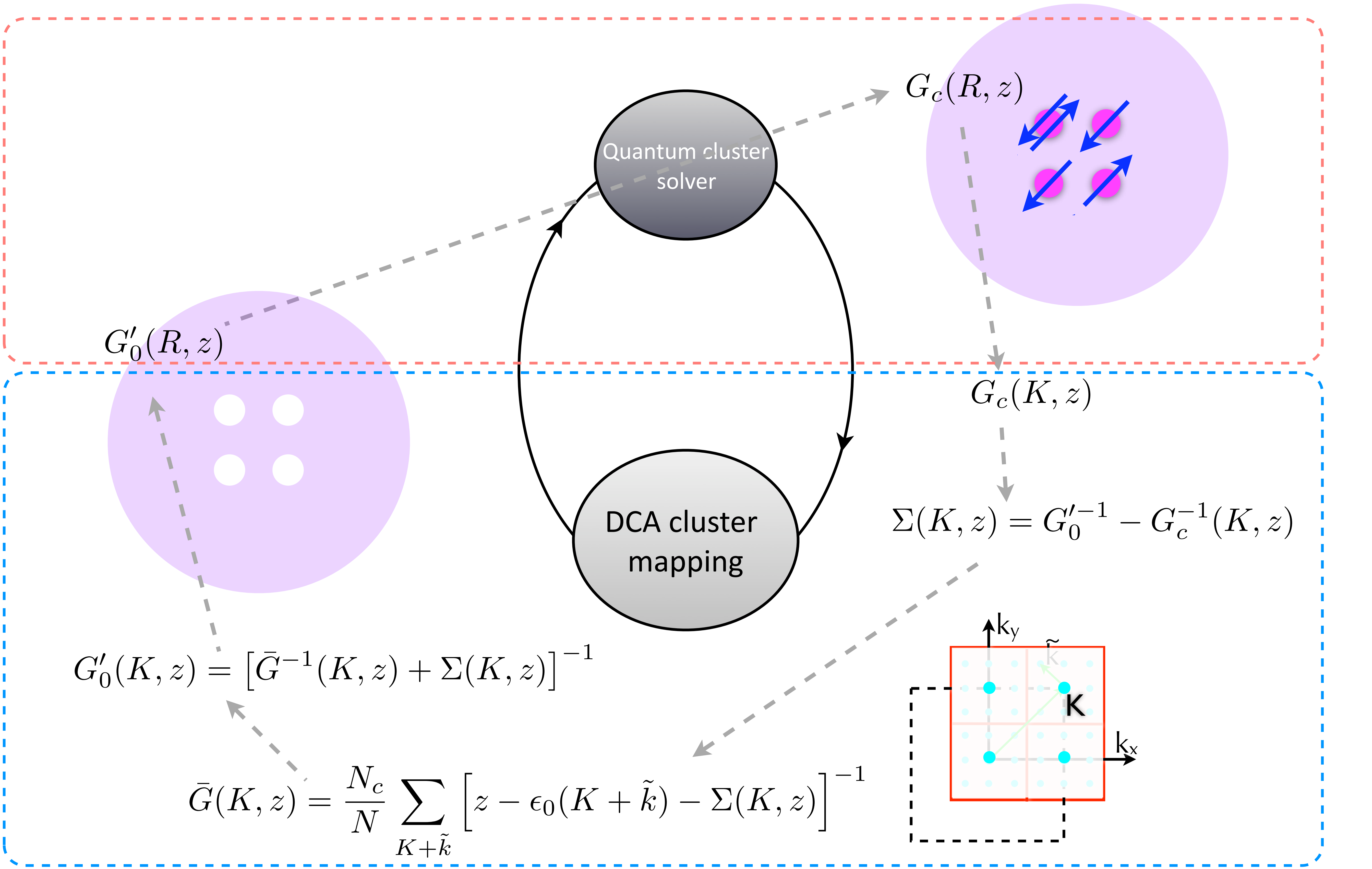
Sketch of the Dynamical Cluster Approximation



Solve many-body problem with quantum Monte Carole on cluster

➢Essential assumption: Correlations are short ranged

DCA method: self-consistently determine the "effective" medium

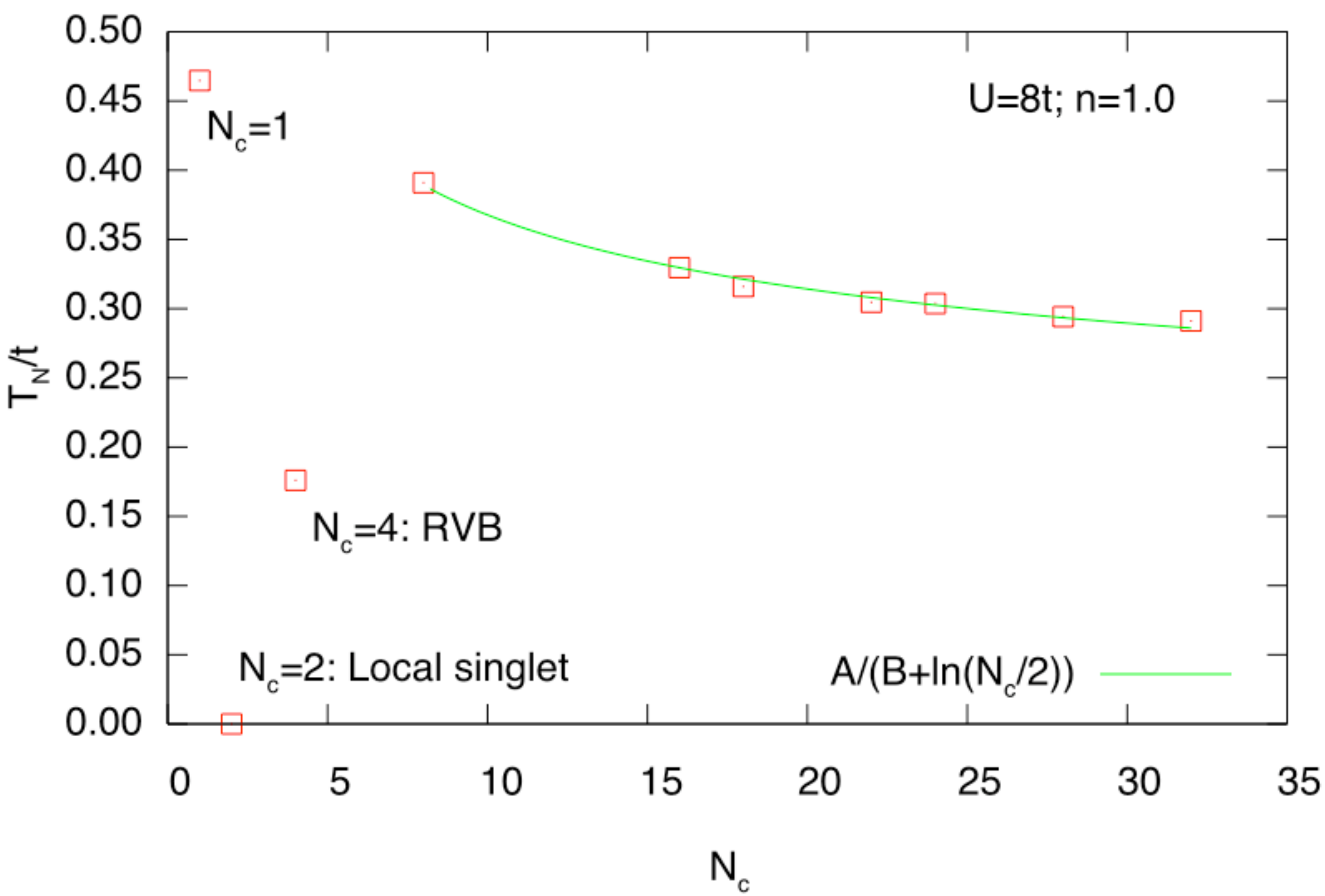


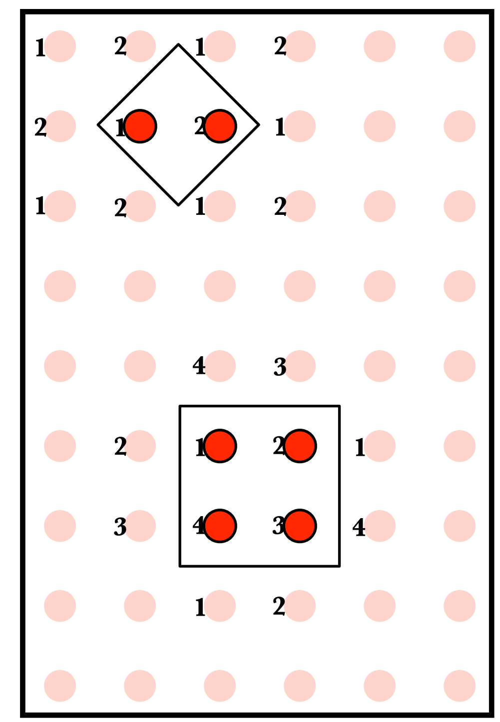
Outline

- Brief introduction into superconductivity and the cuprates
- Background: The two dimensional Hubbard model and the DCA/QMC method
- Simulational studies with the DCA/QMC method
- Algorithmic improvements and a method to study effects of disorder an nanoscale inhomogeneities
  - Accelerating Hirsch-Fye QMC with delayed updates
  - Mixed precision and multithreaded implementations (GPU in particular)
  - Disorder averaging and a first study of how disorder affects the superconducting transition temperature
- DCA++, concurrency, scaling, and performance
  - Results for Cray XT4 and first results for a PF/s scale system
- Summary and conclusions

Cluster size dependence of Néel temperature

No antiferromagnetic order in 2D
Néel temperature indeed vanishes logarithmically with cluster size
(Mermin Wagner Theorem satisfied)





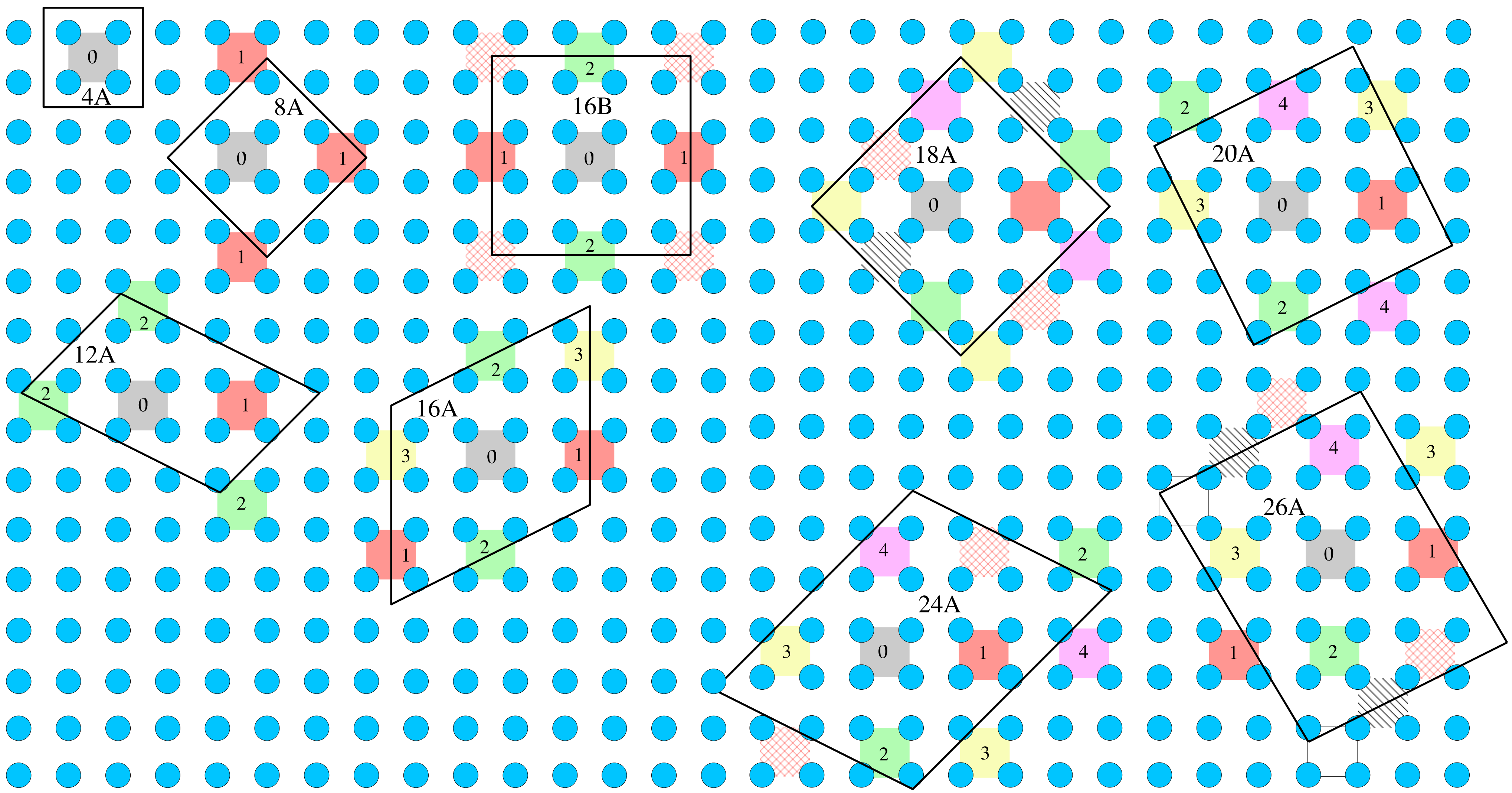
Nc=2:
1 neighbor

Nc=4:
2 neighbors

Maier, Jarrell, Schulthess, Kent, and White, Phys. Rev. Lett. **95**, 237001 (2005)

Simulate larger clusters: Computational tour de force in 2004/2005 on Cray X1E @ NCCS

- Betts *et al.*, for 2D Heisenberg model: (Betts, Can. J. Phys. '99)
  - Selection criteria: symmetry, squareness, # of neighbors in a given shell
- Generalized for *d*-wave pairing in 2D Hubbard model:
  - Count number of neighboring independent 4-site *d*-wave plaquettes

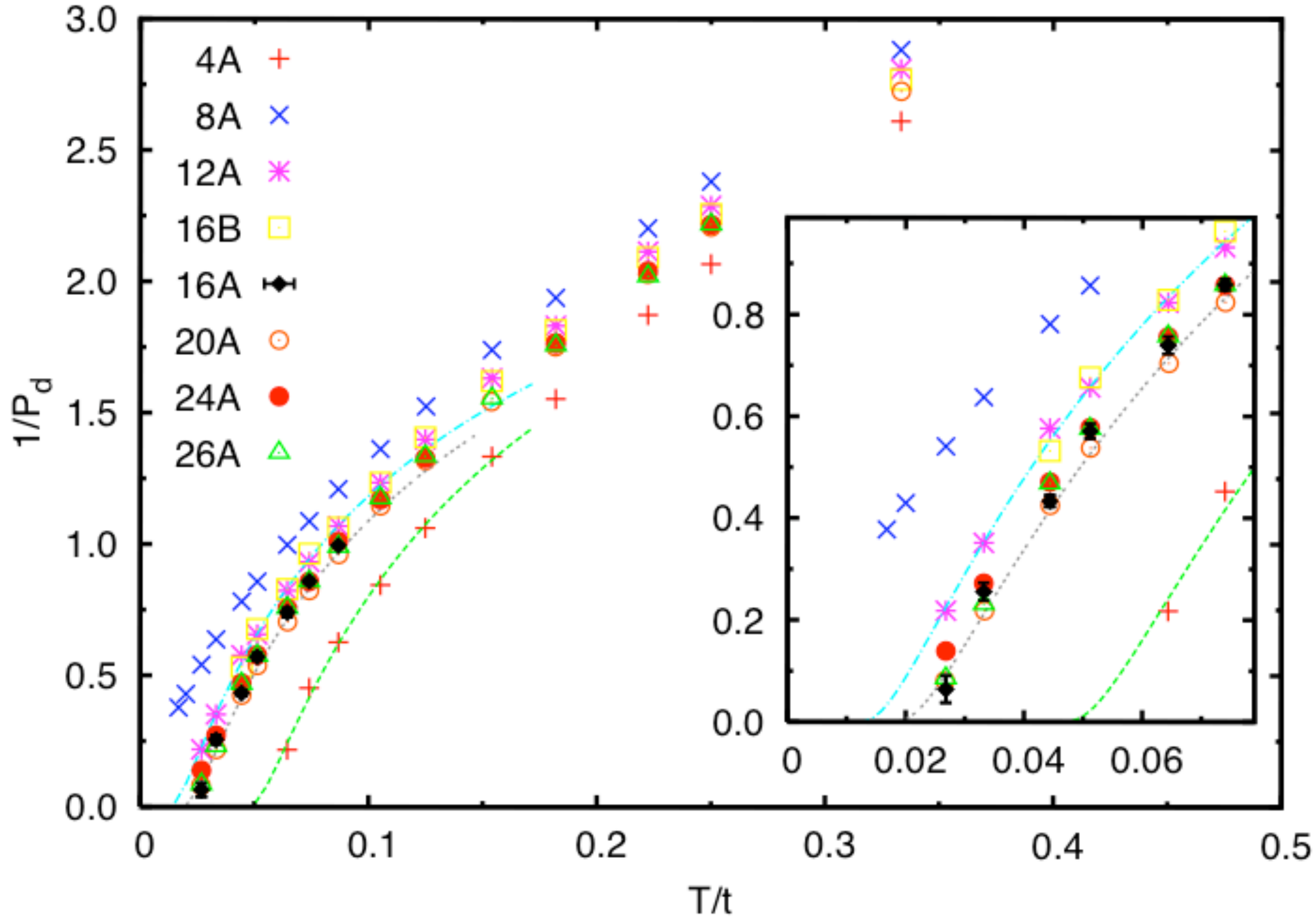


Superconducting transition as function of cluster size: Study divergence of pair-field susceptibility

Measure the pair-field susceptibility $P_d = \int_0^\beta d\tau \langle \Delta_d(\tau)\Delta_d^\dagger(0)\rangle$

$$\Delta_d^\dagger = \frac{1}{2\sqrt{N}} \sum_{l,\delta} (-1)^\delta c_{l\uparrow}^\dagger c_{l+\delta\downarrow}^\dagger$$

Inverse pair-field susceptiblity



Cluster	Z_d
0(MF)	
8A	1
12A	2
16B	2
16A	3
20A	4
24A	4
26A	4

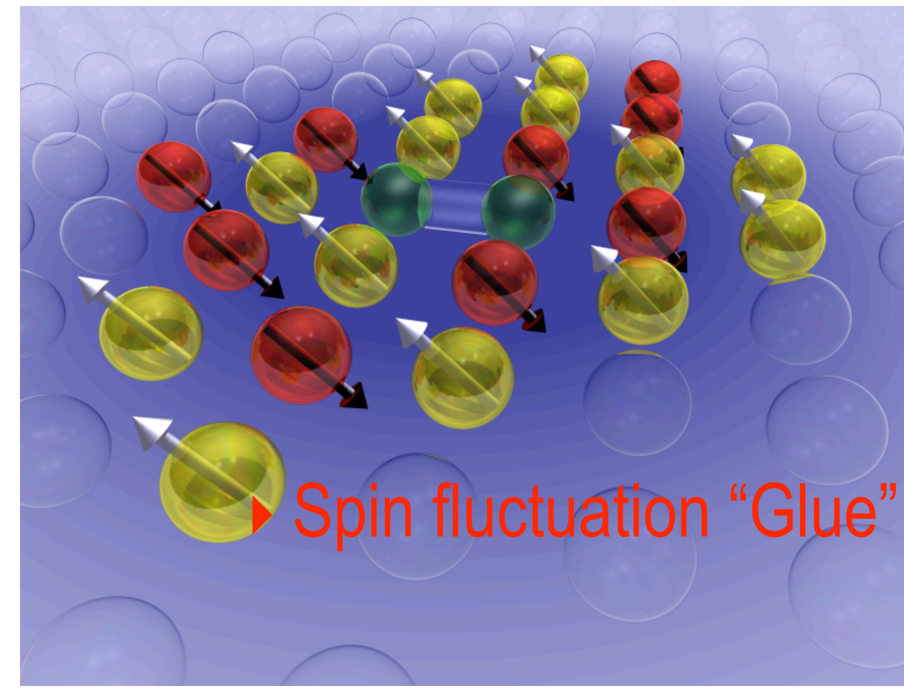
$T_c \approx 0.025t$

Second neighbor shell difficult due to QMC sign problem

Maier, Jarrell, Schulthess, Kent, and White, Phys. Rev. Lett. **95**, 237001 (2005)

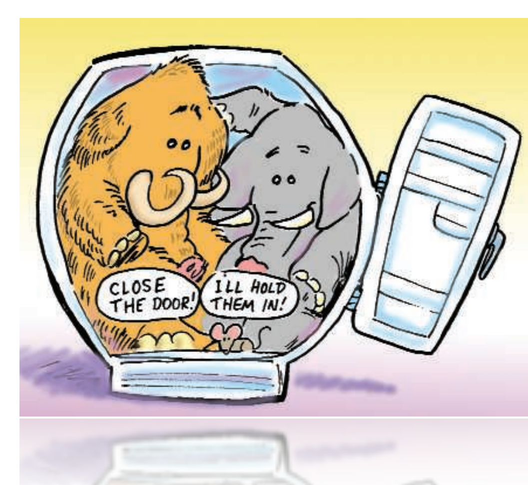
Moving toward a resolution of debate over pairing mechanism in the model

- First systematic solution demonstrates existence of a superconducting transition in 2D Hubbard model   Maier,et al., Phys. Rev. Lett. **95**, 237001 (2005)



- Study the mechanism responsible for pairing in the model
  - Analyze the particle-particle vertex
  - Pairing is mediated by spin fluctuations
    Maier, et al., Phys. Rev. Lett. **96** 47005 (2006)
- Spin susceptibility representation of pairing interactions $3/2\bar{U}^2\chi(q,\omega)$
  - test this for of pairing interaction with neutron scattering and ARPES measurments
  - Maier et al., Phys. Rev. B **75**, 134519 (2007); *ibid* **75**, 144516 (2007)
- Relative importance of spin-fluctuations and resonant valence bond mechanism
  - Maier et al., Phys. Rev. Lett. **100** 237001 (2008)

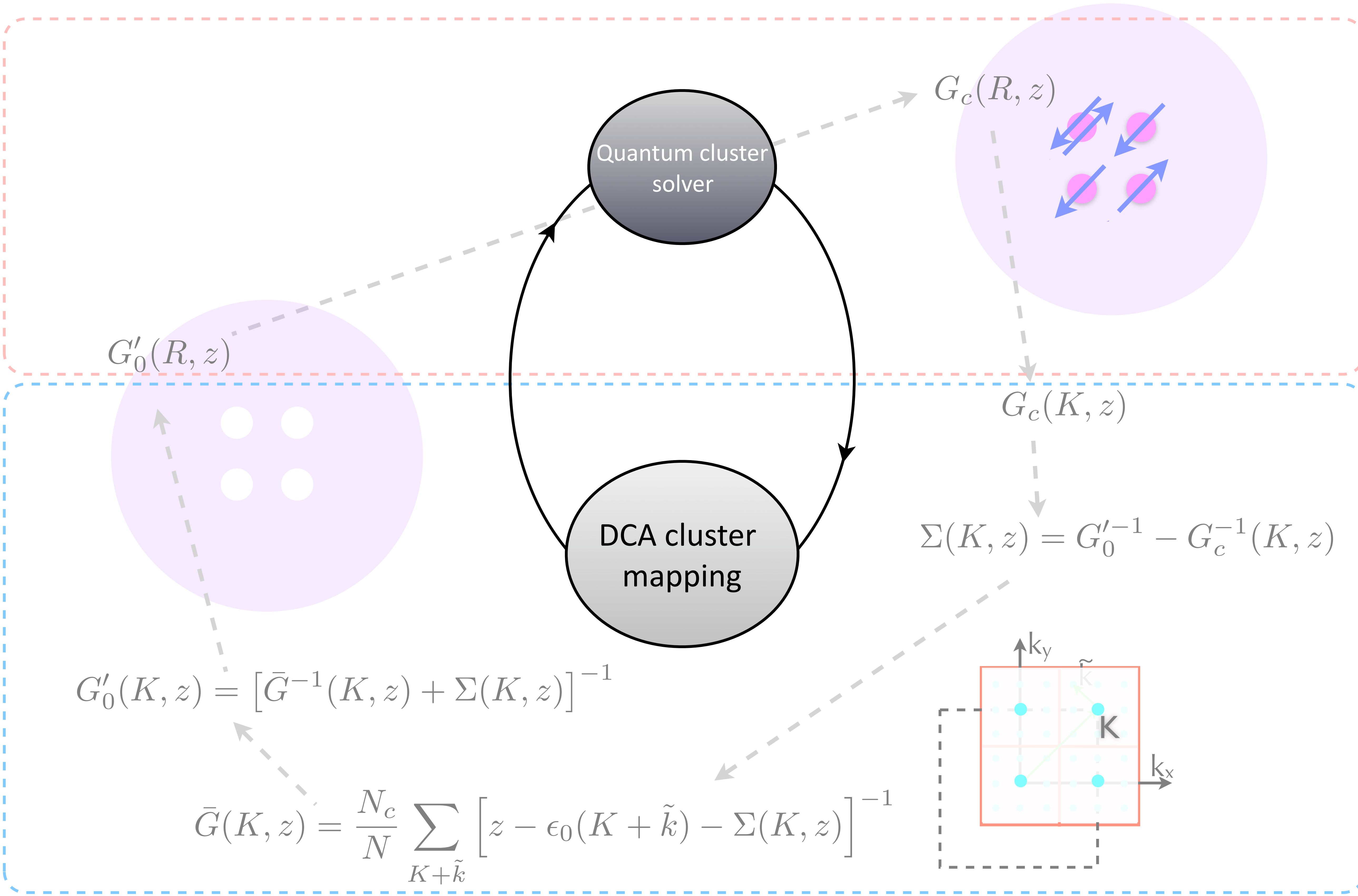
P.W. Anderson, Science **316**, 1705 (2007):
"We have a mammoth (U) and an elephant (J) in our refrigerator - do we care much if there is also a mouse?"
see also http://www.sciencemag.org/cgi/eletters/316/5832/1705
"Scalapino is not a glue-sniffer"



Outline

- Brief introduction into superconductivity and the cuprates
- Background: The two dimensional Hubbard model and the DCA/QMC method
- Simulational studies with the DCA/QMC method
- Algorithmic improvements and a method to study effects of disorder an nanoscale inhomogeneities
  - Accelerating Hirsch-Fye QMC with delayed updates
  - Mixed precision and multithreaded implementations (GPU in particular)
  - Disorder averaging and a first study of how disorder affects the superconducting transition temperature
- DCA++, concurrency, scaling, and performance
  - Results for Cray XT4 and first results for a PF/s scale system
- Summary and conclusions

DCA method: self-consistently determine the "effective" medium



Hirsch-Fye Quantum Monte Carole (HF-QMC) for the quantum cluster solver

Hirsch & Fye, Phys. Rev. Lett. **56**, 2521 (1998)

Partition function & Metropolis Monte Carlo $Z = \int e^{-E[\mathbf{x}]/k_\mathrm{B}T} d\mathbf{x}$

Acceptance criterion for M-MC move: $\min\{1, e^{E[\mathbf{x}_k]-E[\mathbf{x}_{k+1}]}\}$

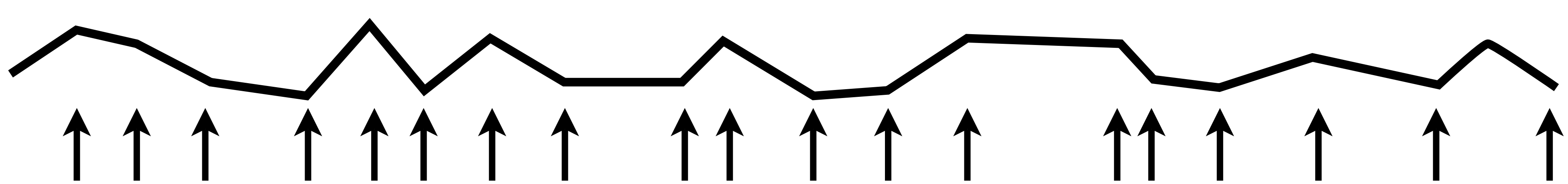
Partition function & HF-QMC: $Z \sim \sum_{s_i,l} \det[\mathbf{G}_c(s_i,l)^{-1}]$

N_c $N_l \approx 10^2$

$N_t = N_c \times N_l \approx 2000$

matrix of dimensions $N_t \times N_t$

Acceptance: $\min\{1, \det[\mathbf{G}_c(\{s_i,l\}_k)]/\det[\mathbf{G}_c(\{s_i,l\}_{k+1})]\}$

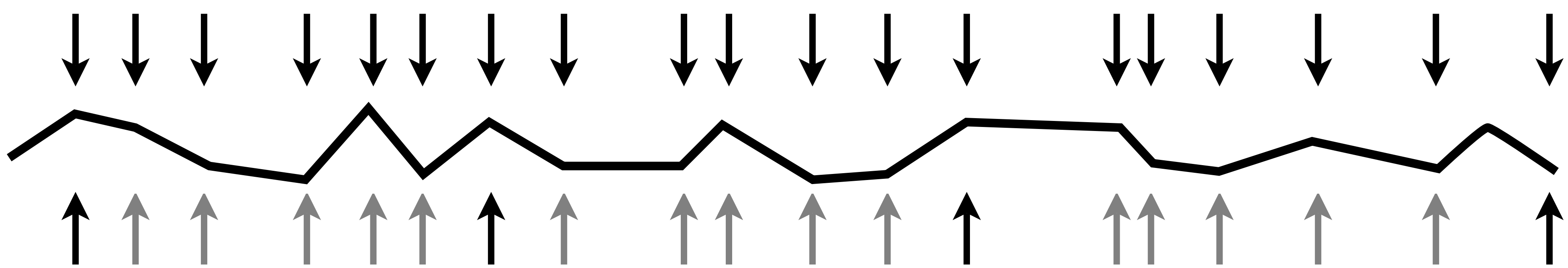


Update of accepted Green's function:

$$\mathbf{G}_c(\{s_i,l\}_{k+1}) = \mathbf{G}_c(\{s_i,l\}_k) + \mathbf{a}_k \times \mathbf{b}_k$$

HF-QMC with Delayed updates (or Ed updates)

$$\mathbf{G}_c(\{s_i,l\}_{k+1}) = \mathbf{G}_c(\{s_i,l\}_k) + \mathbf{a}_k \times \mathbf{b}_k^t$$

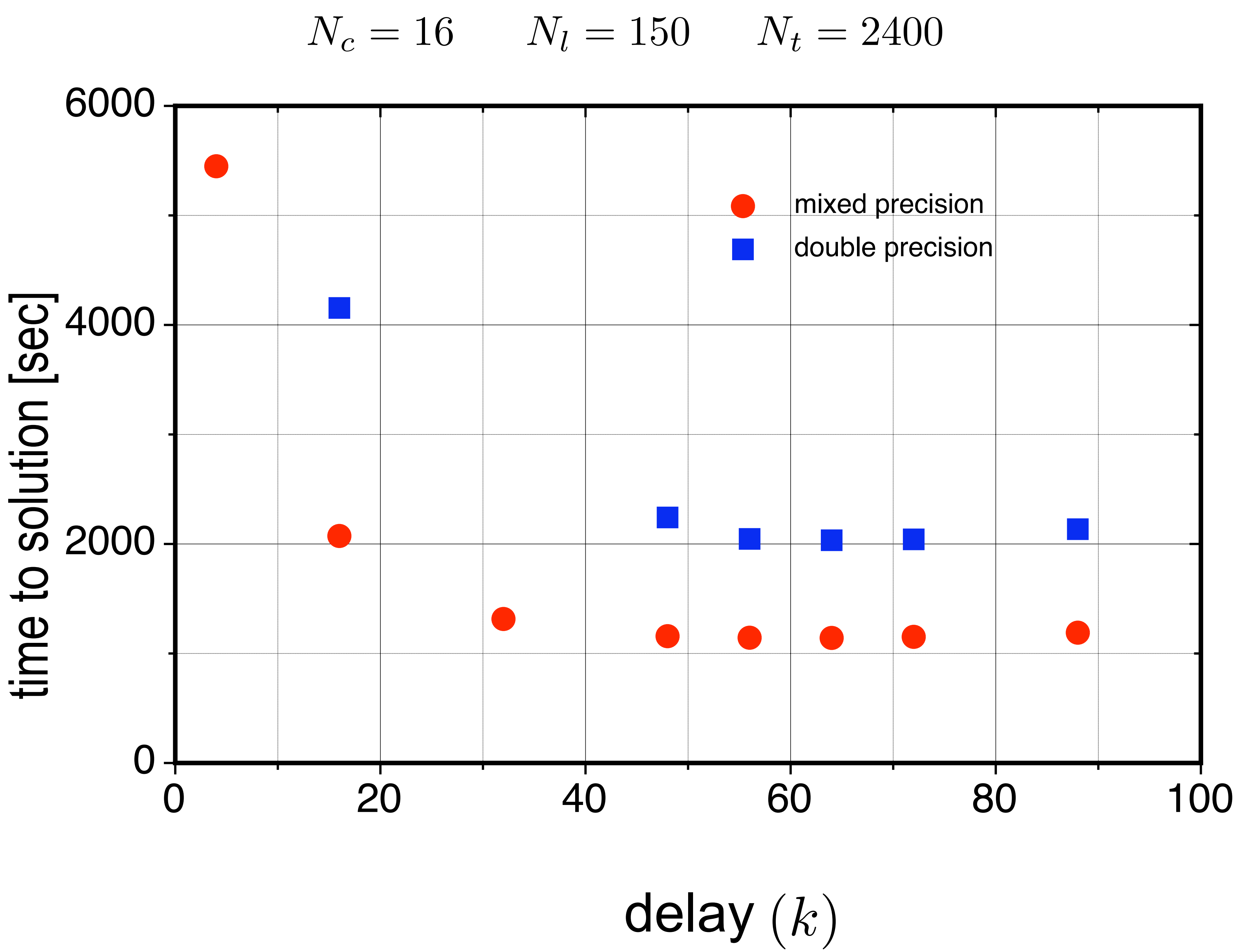


$$\mathbf{G}_c(\{s_i,l\}_{k+1}) = \mathbf{G}_c(\{s_i,l\}_0) + [\mathbf{a}_0|\mathbf{a}_1|...|\mathbf{a}_k] \times [\mathbf{b}_0|\mathbf{b}_1|...|\mathbf{b}_k]^t$$

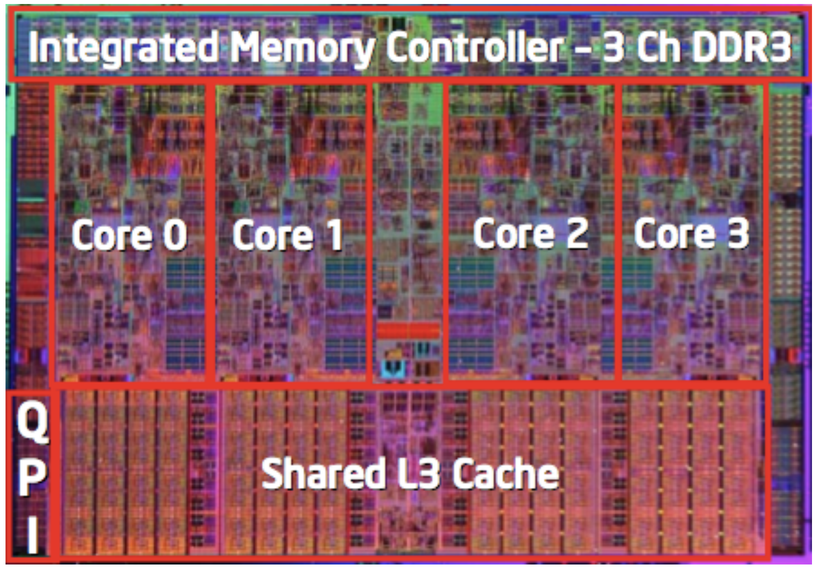
Complexity for *k* updates remains $\mathcal{O}(kN_t^2)$

But we can replace *k* rank-1 updates with one matrix-matrix multiply plus some additional bookkeeping.

Performance improvement with delayed updates

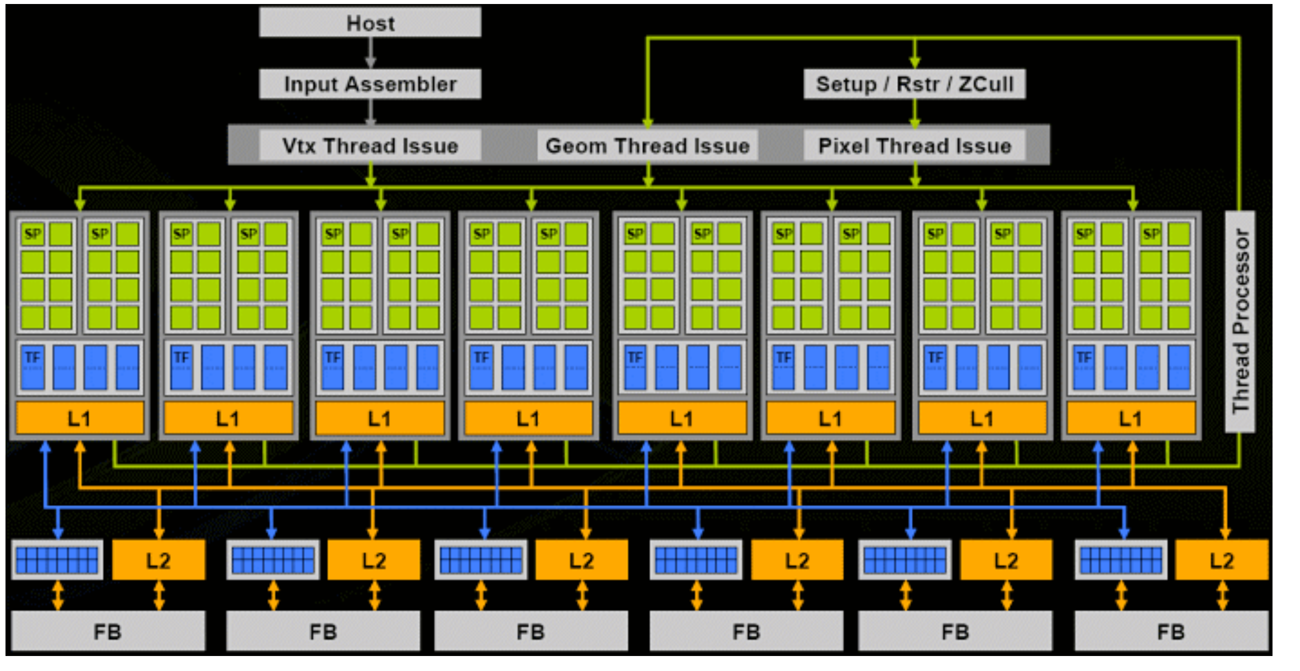


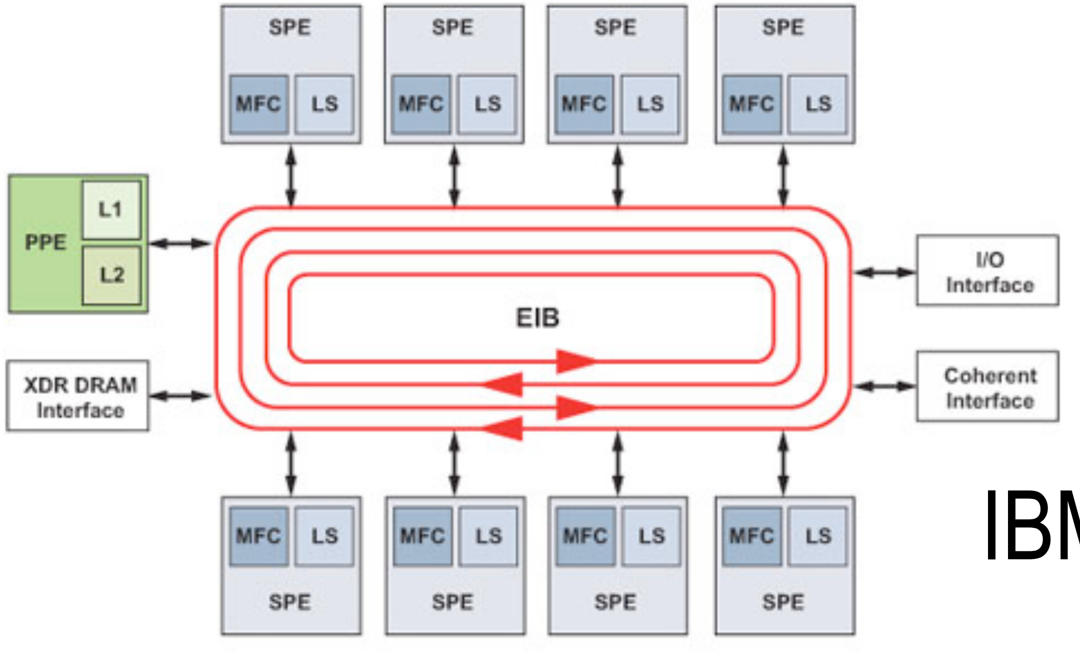
MultiCore/GPU/Cell: threaded programming



Multi-core processors: OpenMP (or just MPI)

NVIDIA G80 GPU: CUDA, cuBLAS

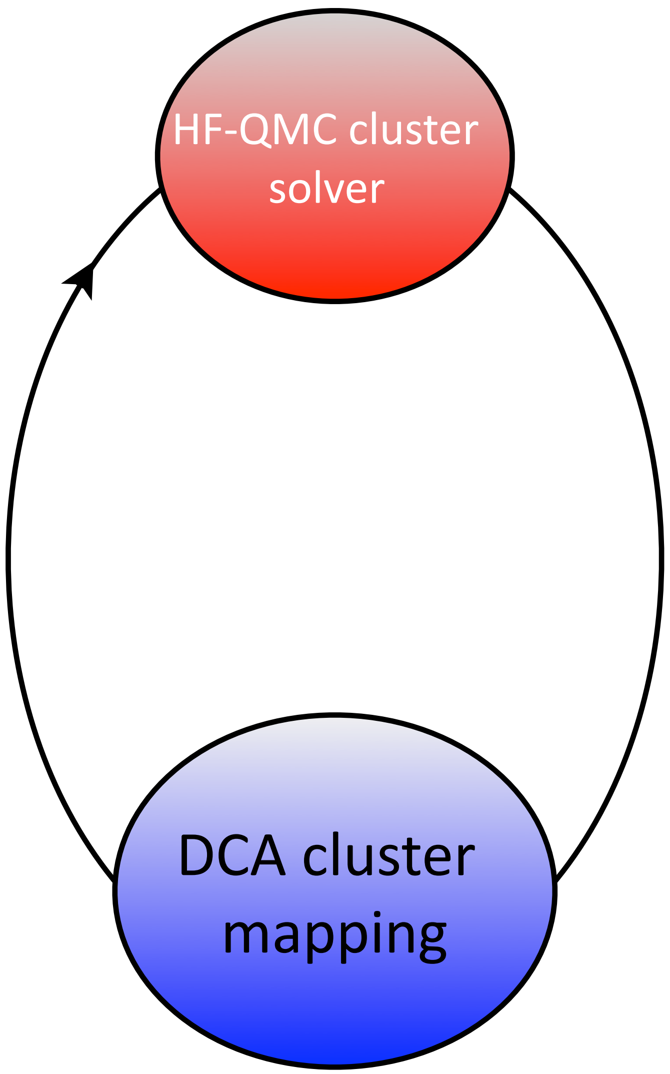




IBM Cell BE: SIMD, threaded prog.

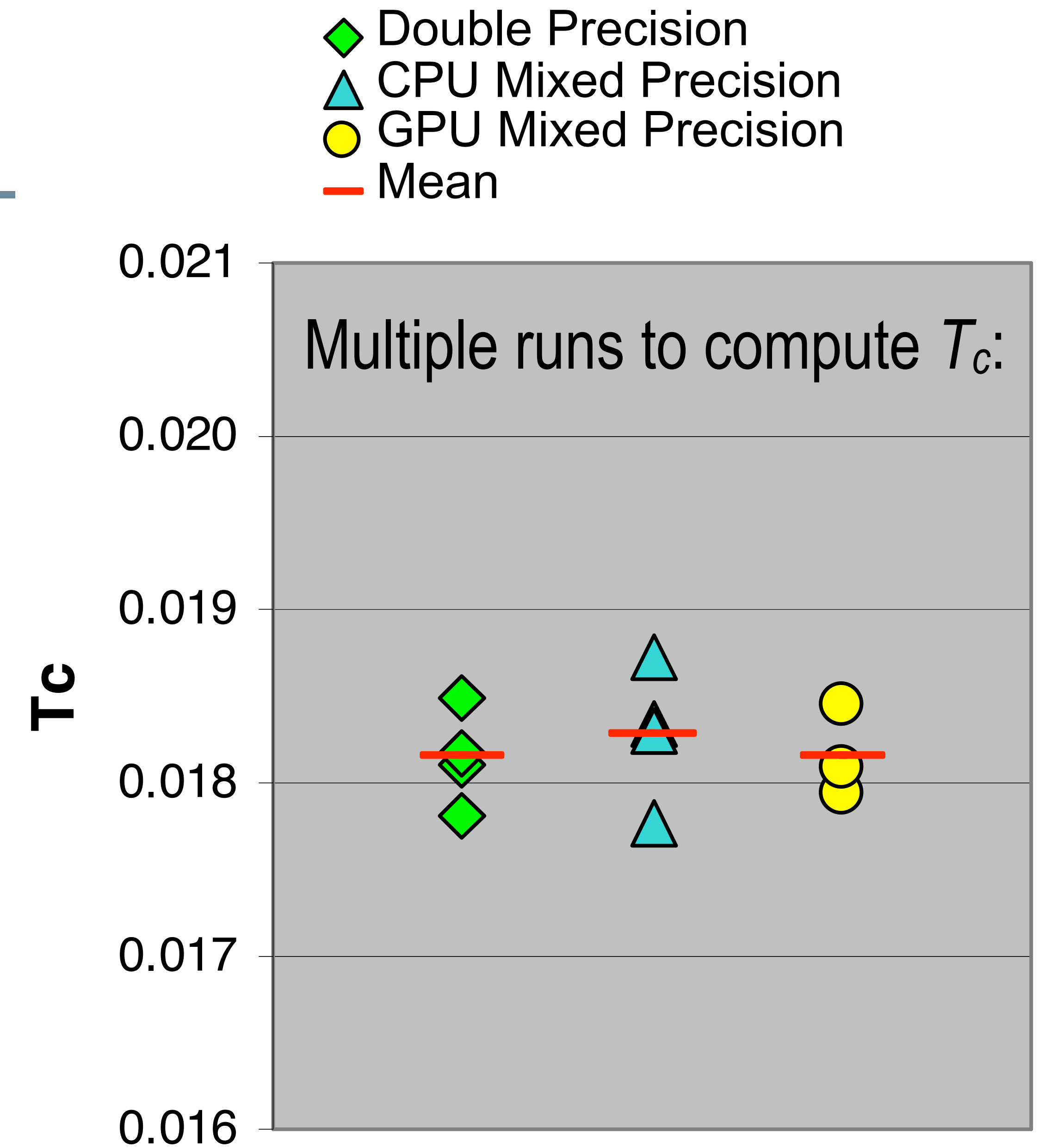
DCA++ with mixed precision

Run HF-QMC in single precision



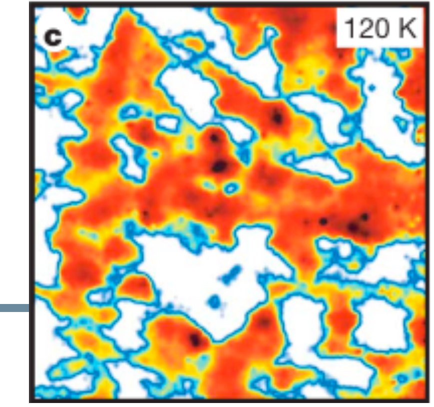
Keep the rest of the code, in particular cluster mapping in double precision

Results for mixed and double precision runs are identical for same random number sequence!



Speedup of HF-QMC updates (2GHz Opteron vs. NVIDIA 8800GTS GPU):
- 9x for offloading BLAS to GPU & transferring all data
- 13x for offloading BLAS to GPU & lazy data transfer
- 19x for full offload HF-updates & full lazy data transfer

Disorder and inhomogeneities



Hubbard Model with random disorder (eg. in U)

$$H^{(\nu)} = -t \sum_{\langle ij\rangle,\sigma} c^\dagger_{i\sigma} c_{j\sigma} + \sum_i U_i^{(\nu)} n_{i\uparrow} n_{i\downarrow}$$

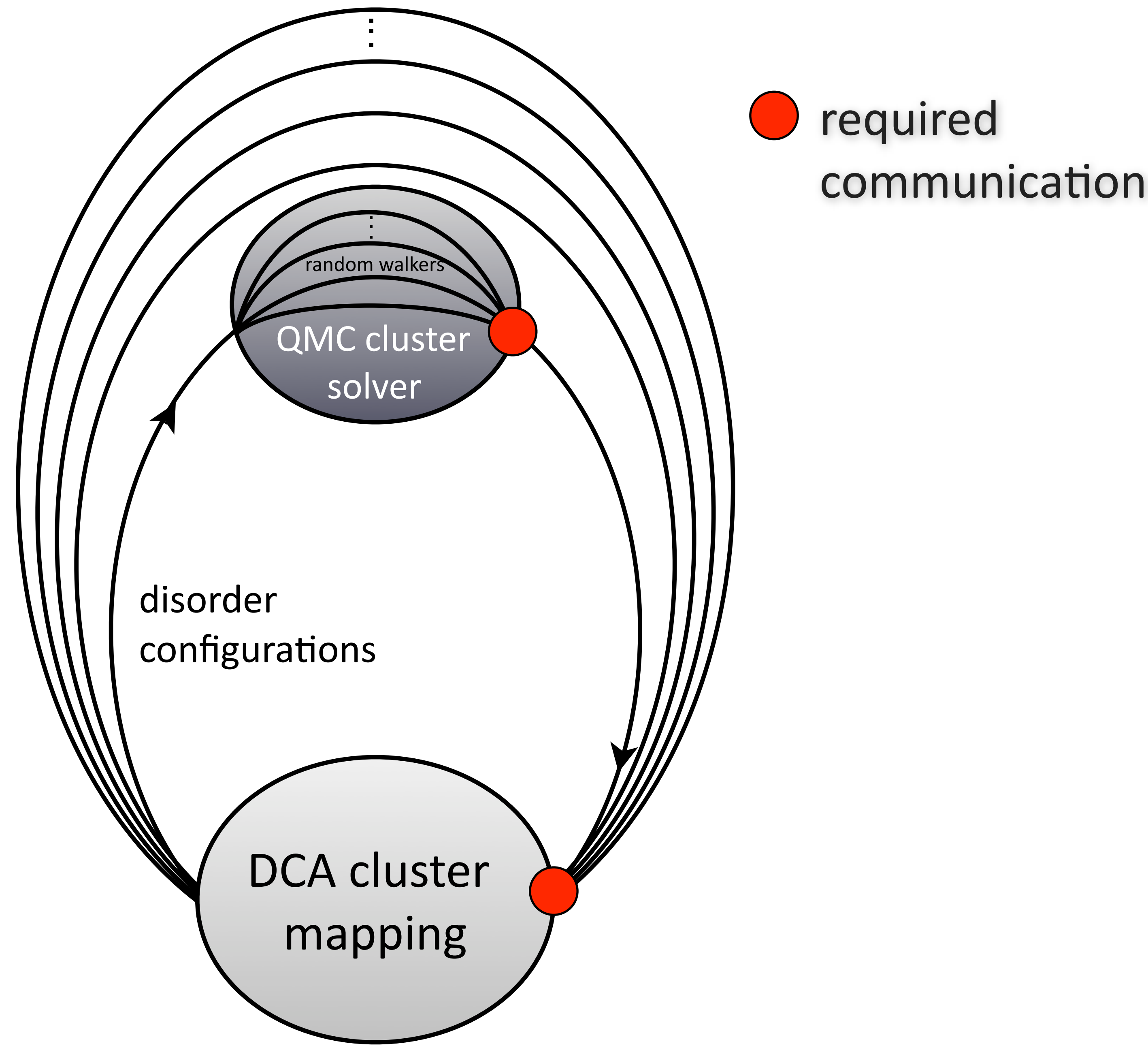
$$U_i^{(\nu)} \in \{U, 0\}; N_c = 16 \rightarrow N_d = 2^{16}$$

... need to disorder-average cluster Green function

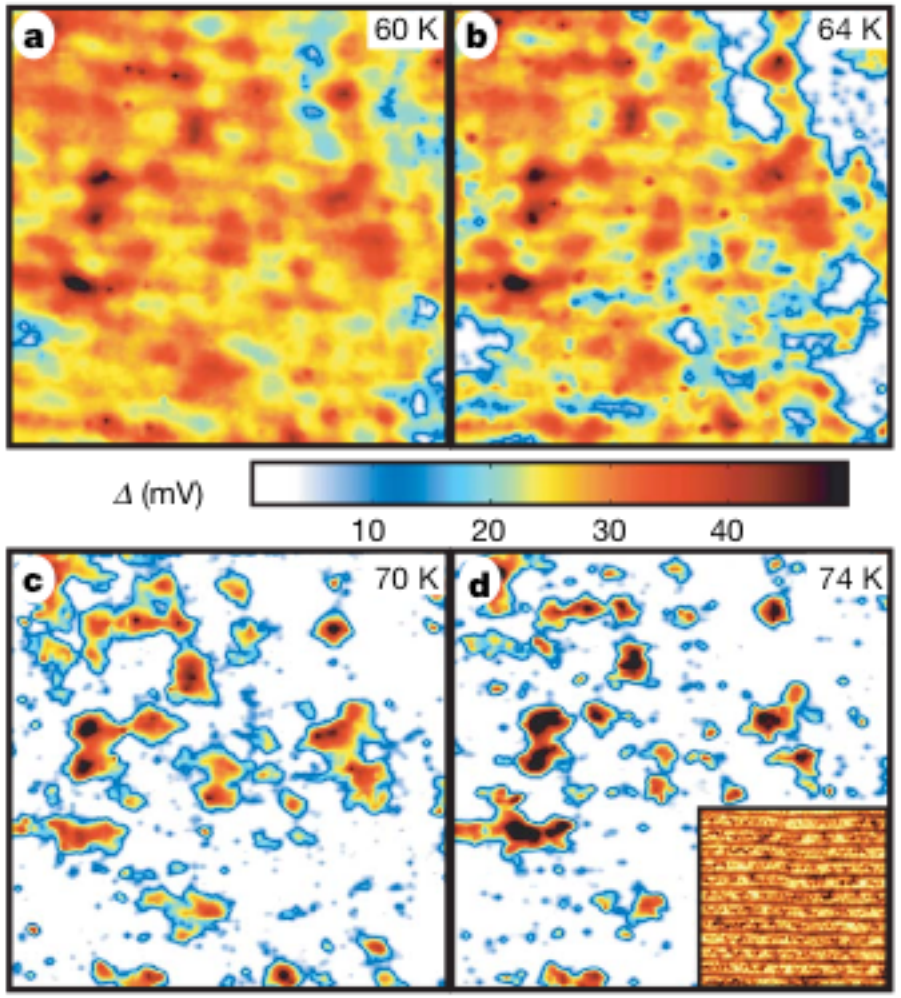
$$G_c(X_i - X_j, z) = \frac{1}{N_c} \sum_{\nu=1}^{N_d} G_c^\nu(X_i, X_j, z)$$

Algorithm 1 DCA/QMC Algorithm with QMC cluster solver (lines 5-10), disorder averaging (lines 4, 11-12), and DCA cluster mapping (line 3, 13)

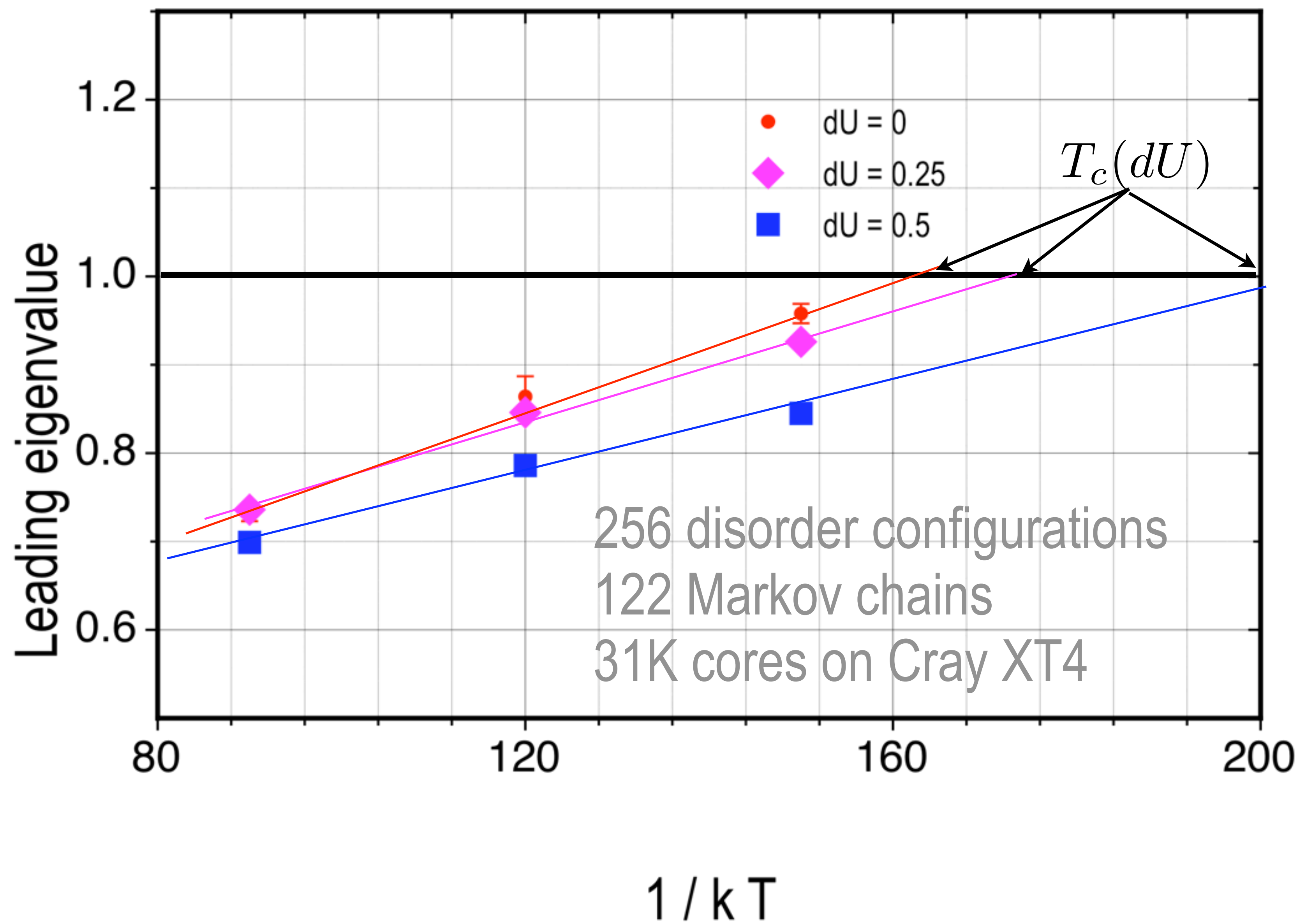
1: Set initial self-energy
2: **repeat**
3: Compute the coarse-grained Green Function
4: **for** Every disorder configuration (in parallel) **do**
5: Perform warm-up steps
6: **for** Every Markov chain (in parallel) **do**
7: Update auxiliary fields
8: Measure Green Function and observables
9: **end for**
10: Accumulate measurements over Markov chains
11: **end for**
12: Accumulate measurements over disorder configurations.
13: Re-compute the self-energy
14: **until** self consistency is reached



Effect of disorder in U $U_i = U \pm \nu_i dU \quad P_d = \frac{P_d^0}{1 - \Gamma^{pp} P_d^0}$



Temperature evolution of the superconducting gap taken on a 300 A area of a cuprate with T_c=65K [reproduced from Gomez et al. Nature **447**, 569-572 (2007)]. The gap varies spatially on a scale of 1-3 nm and persists in some regions to temperatures well above T_c as can be seen from panel c and d.

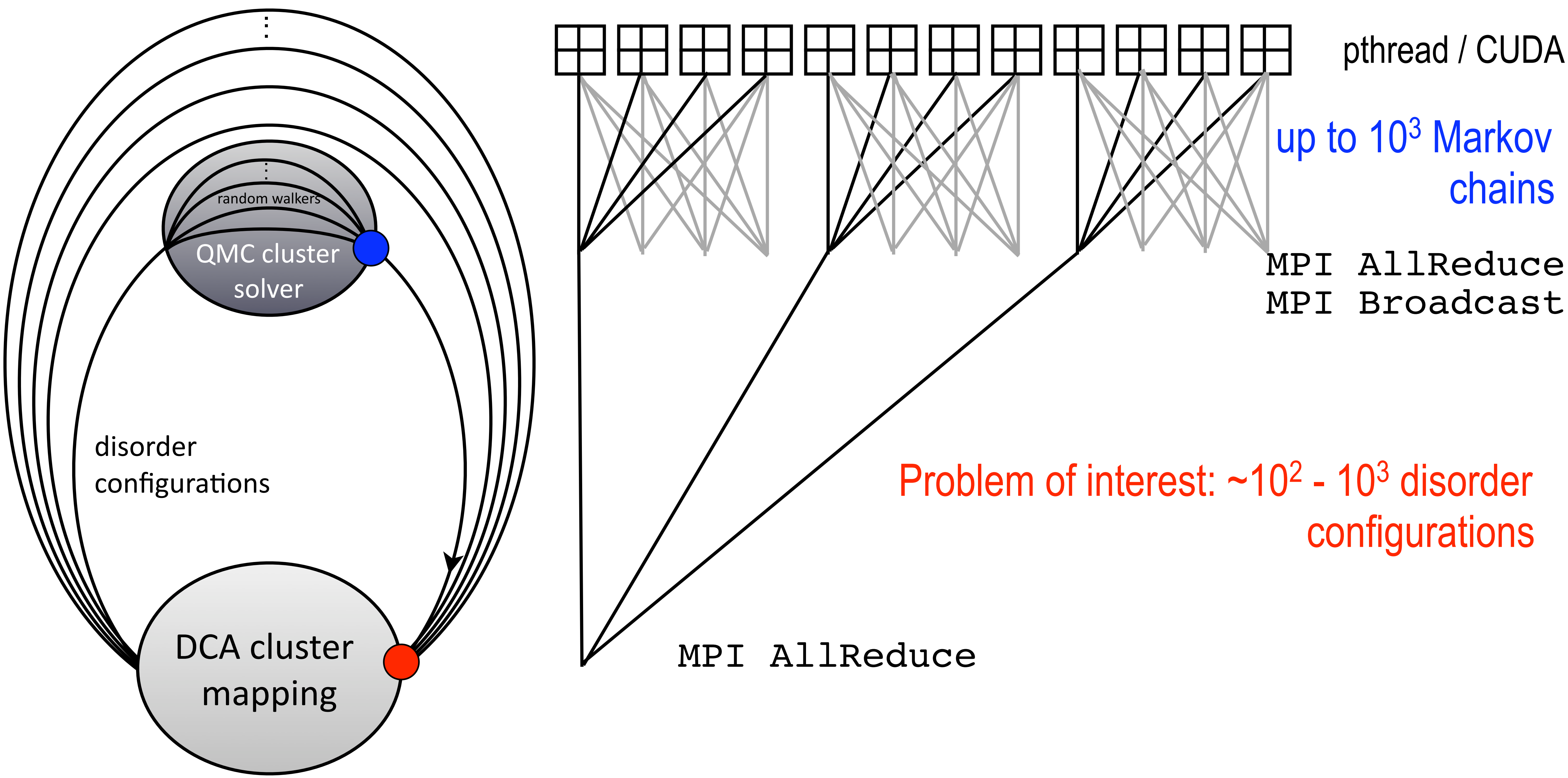


Disorder reduced transition temperature
- how does it affect pairing strength?
- does to pairing strength vary spatially?
- what about other types of disorder?
- relationship to chemistry of materials?
- .....

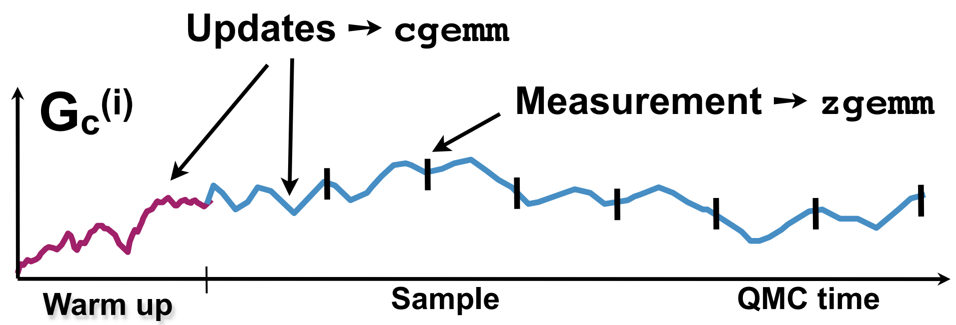
Outline

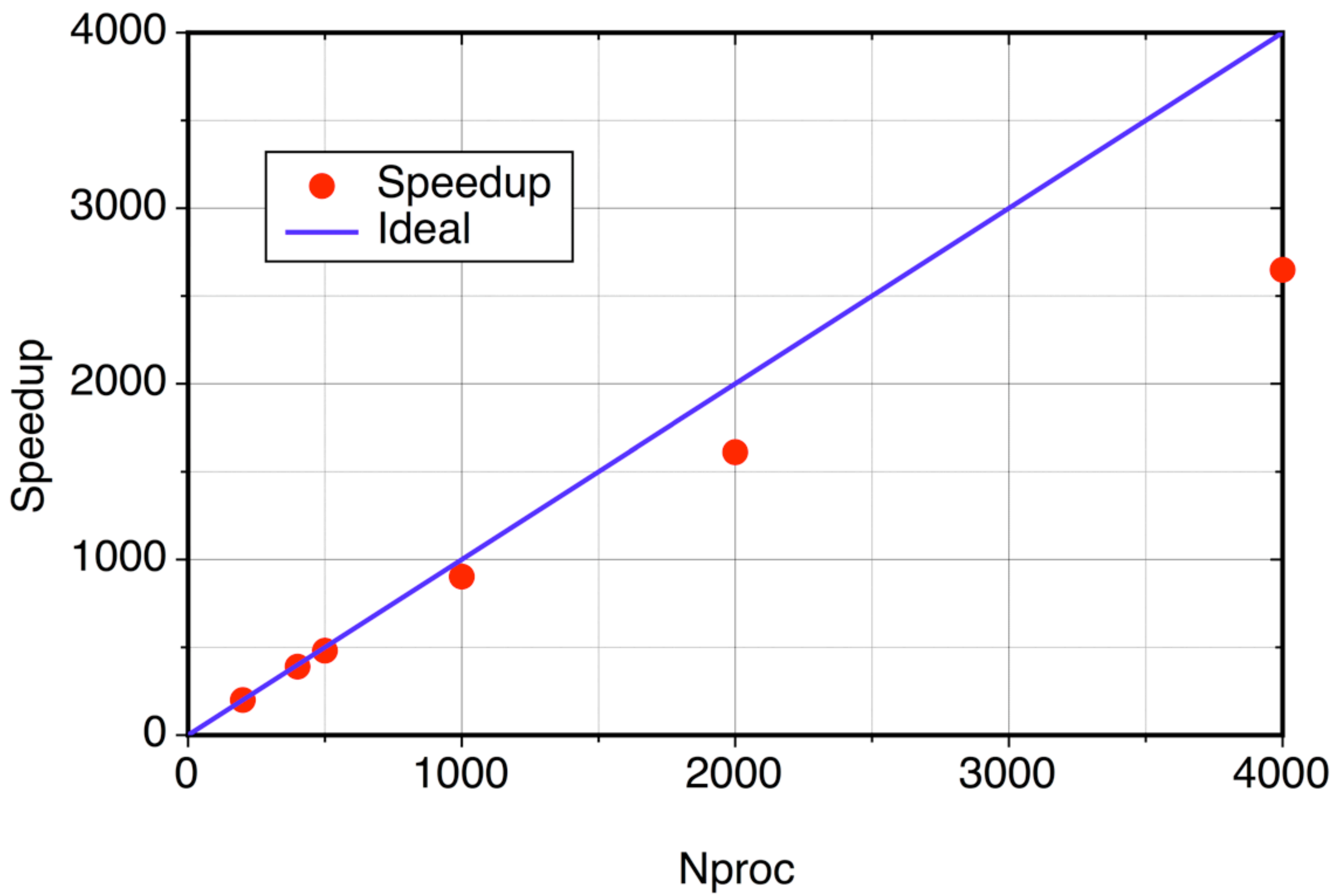
- Brief introduction into superconductivity and the cuprates
- Background: The two dimensional Hubbard model and the DCA/QMC method
- Simulational studies with the DCA/QMC method
- Algorithmic improvements and a method to study effects of disorder an nanoscale inhomogeneities
  - Accelerating Hirsch-Fye QMC with delayed updates
  - Mixed precision and multithreaded implementations (GPU in particular)
  - Disorder averaging and a first study of how disorder affects the superconducting transition temperature
- DCA++, concurrency, scaling, and performance
  - Results for Cray XT4 and first results for a PF/s scale system
- Summary and conclusions

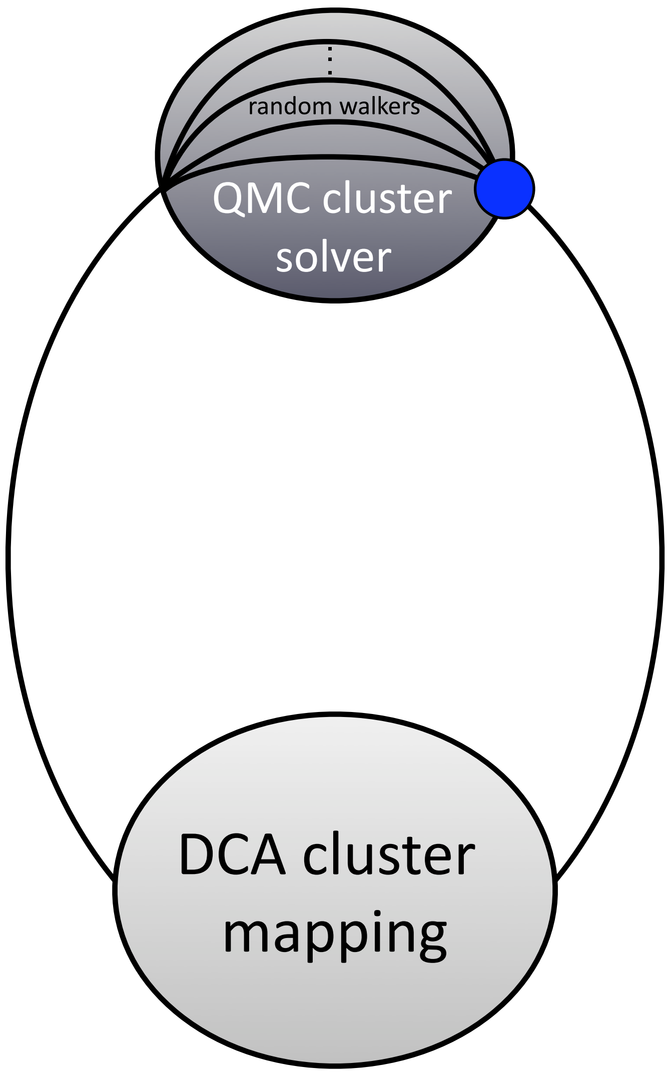
DCA++ code from a concurrency point of view



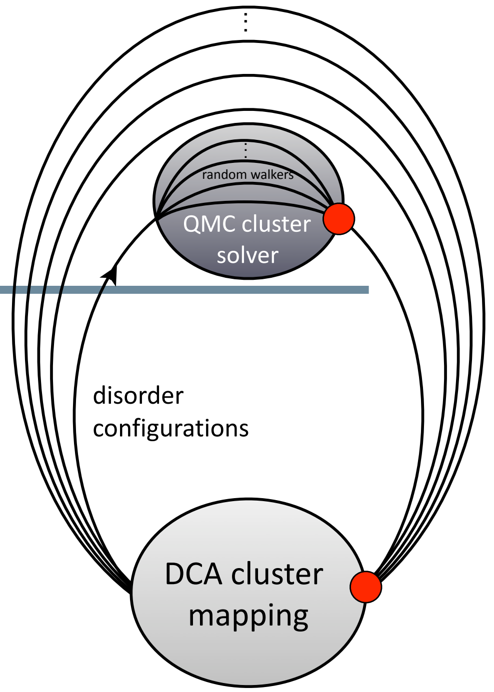
DCA++: strong scaling on HF-QMC



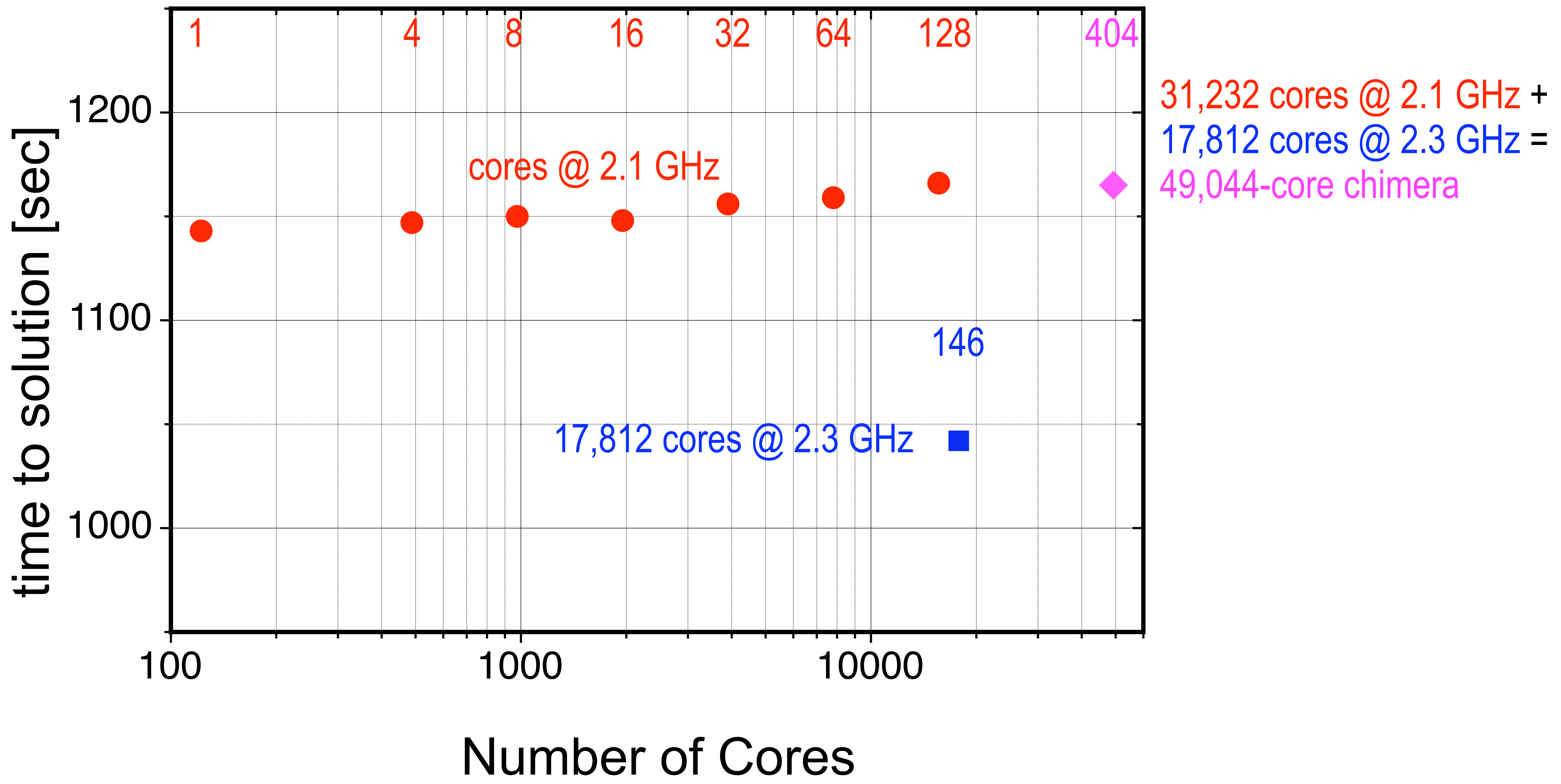




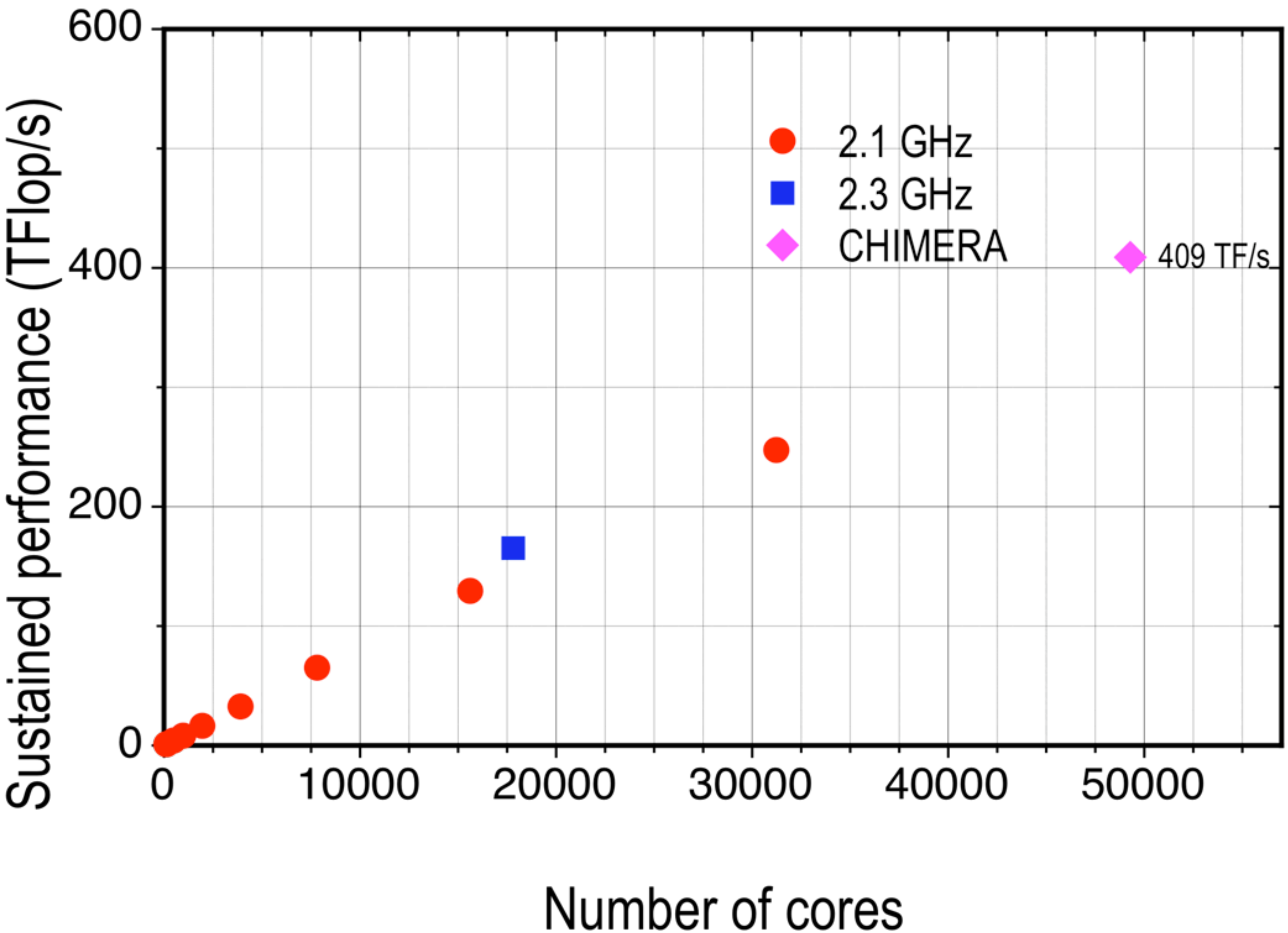
Weak scaling on Cray XT4



- HF-QMC: 122 Markov chains on 122 cores
- Weak scaling over disorder configurations



Sustained performance of DCA++ on Cray XT4

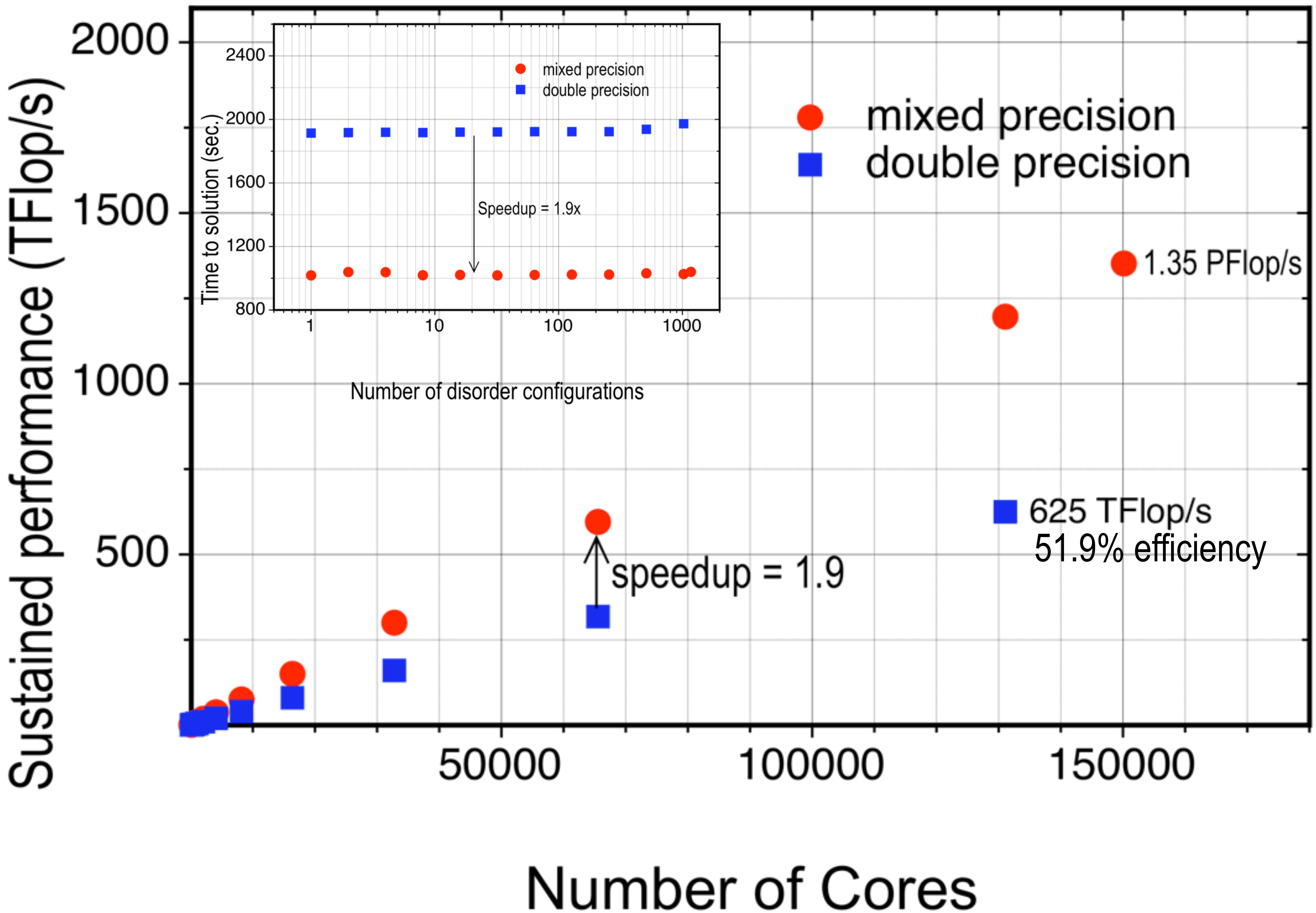


Cray XT5 portion of Jaguar @ NCCS

Peak: 1.382 TF/s
Quad-Core AMD
Freq.: 2.3 GHz
150,176 cores
Memory: 300 TB
For more details, go to
www.nccs.gov

Sustained performance of DCA++ on Cray XT5

Weak scaling with number disorder configurations, each running on 128 Markov chains on 128 cores (16 nodes) - 16 site cluster and 150 time slides



Outline

- Brief introduction into superconductivity and the cuprates
- Background: The two dimensional Hubbard model and the DCA/QMC method
- Simulational studies with the DCA/QMC method
- Algorithmic improvements and a method to study effects of disorder an nanoscale inhomogeneities
  - Accelerating Hirsch-Fye QMC with delayed updates
  - Mixed precision and multithreaded implementations (GPU in particular)
  - Disorder averaging and a first study of how disorder affects the superconducting transition temperature
- DCA++, concurrency, scaling, and performance
  - Results for Cray XT4 and first results for a PF/s scale system
- Summary and conclusions

Summary

- Today's methods and computational capabilities allow us to take a deep look into the mechanisms of high-T_c superconductivity
  - Simulations of superconducting transition in model without phonons
  - Dominant contribution to pairing mechanism: "glue" due to spin fluctuations
- DCA++ - optimally mapping DCA/QMC method onto today's hardware architectures
  - Algorithm: Hirsch-Fye QMC with delayed updates (>10x speedup)
  - Accelerator work motivated: mixed precision (almost 2x speedup)
  - Highly scalable implementation to study disorder and nanoscale inhomegeneities
  - Extensible implementation based on C++/STL generic programming model
- Sustained 1.35 PF/s on 150K cores of Cray XT5 portion of NCCS/Jaguar
  - Sustained 625 TF/s on 130K cores in double precision (52% efficiency)
- More than 1000 fold capability enhancement since 2004:
  - NCCS 2004: Cray X1E with 18TF/s peak, DCA/QMC sustained about 8TF/s (required high memory bandwidth)
  - NCCS 2008: factor 50-100 more in peak Flop/s & at least 20x due to algorithms
  - Future: Continuous time QMC - a new class of QMC algorithms

HPC in the age of massively parallel processing (MPP) architectures: what does this really mean?

Evolution of the fastest sustained performance in real simulations

~1 Exaflop/s
~10^7 processing units

1.35 Petaflop/s
Cray XT5
1.5 10^5 processor cores

1.02 Teraflop/s
Cray T_{3D}
1.5 10^3 processors

1.5 Gigaflop/s
Cray YMP
0.8 10^1 processors

1989 1998 2008 2018

New algorithm to enable 1+ PFlop/s sustained performance in simulations of disorder effects in high-T_c superconductors

G. Alvarez
M. S. Summers
D. E. Maxwell
M. Eisenbach
J. S. Meredith
J. M. Larkin
J. Levesque
T. A. Maier
P. R. C. Kent
E. F. D'Azevedo
T. C. Schulthess

D. Scalapino
M. Jarrell
J. Vetter
Trey White
staff at NCCS & Cray
many others

Computational resources:
NCCS @ ORNL

Funding:
ORNL-LDRD,
DOE-ASCR,
DOE-BES

Oak Ridge National Laboratory

New algorithm to enable 1+ PFlop/s sustained performance in simulations of disorder effects in high-T_c superconductors

Models,
Methods,
& Implementation

Map to Hardware

Operations

System design

T. A. Maier
P. R. C. Kent
T. C. Schulthess
G. Alvarez
M. S. Summers
E. F. D'Azevedo
J. S. Meredith
M. Eisenbach
D. E. Maxwell
J. M. Larkin
J. Levesque

Physics

Software

Comp. mathematics

Computer Science

Computer Center

Hardware vendor

OAK RIDGE
National Laboratory

Conclusions / Challenges

- DCA++ was just one of several successful early application teams (MAD++, CHIMERA, S3D, GTC, GW-LSMS, ...)
  - All were interdisciplinary teams lead by scientists with heavy involvement of Comp. Math, CS, Operations, & Systems/Vendor
- More such teams have to develop outside of ORNL / DOE
  - Scientists are seriously looking at HPC - how can they be engaged?
- These teams need the same mix as the successful ORNL models:
  - Scientists (lead) — find senior scientists who invest/engage
  - Software developers (mostly from science teams) — enough scientists have to understand methods / math. / software development
  - Computational Mathematics — funding agencies have to grow these programs
  - Computer Science (hardware oriented) — CS has to shift focus from software design to hardware systems
  - Operations (NCCS and smaller centers)
  - Hardware / system integrators (vendors) — vendors have to disseminate components early in development phase